

Measuring school demand in the presence of spatial dependence. A conditional approach

Laura López-Torres and Diego Prior

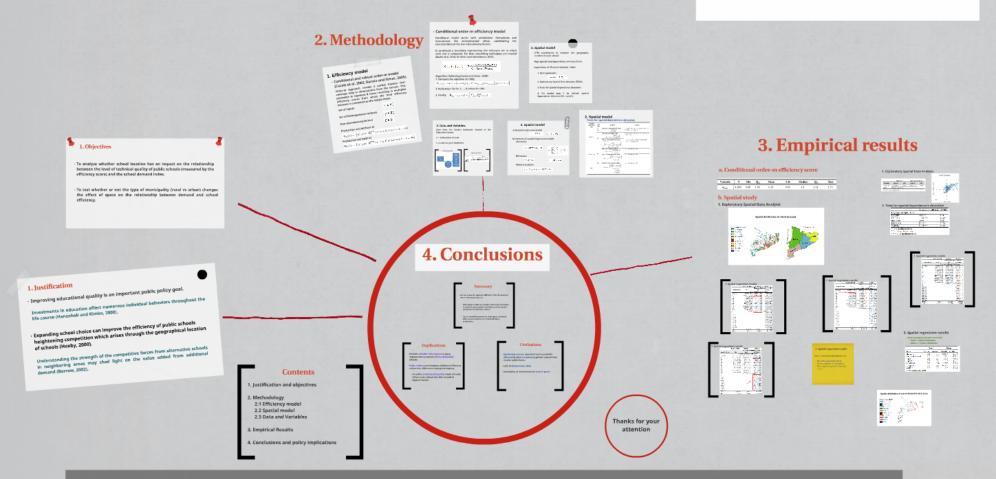












Measuring school demand in the presence of spatial dependence. A conditional approach

Laura López-Torres and Diego Prior

# **Contents**

- 1. Justification and objectives
- 2. Methodology
  - 2.1 Efficiency model
  - 2.2 Spatial model
  - 2.3 Data and Variables
- 3. Empirical Results
- 4. Conclusions and policy implications

### 1. Justification

Improving educational quality is an important public policy goal.

Investments in education affect numerous individual behaviors throughout the life course (Hanushek and Kimko, 2000).

 Expanding school choice can improve the efficiency of public schools heightening competition which arises through the geographical location of schools (Hoxby, 2000).

Understanding the strength of the competitive forces from alternative schools in neighboring areas may shed light on the value added from additional demand (Barrow, 2002).





# 1. Objectives

- To analyze whether school location has an impact on the relationship between the level of technical quality of public schools (measured by the efficiency score) and the school demand index.
- To test whether or not the type of municipality (rural vs urban) changes the effect of space on the relationship between demand and school efficiency.

# 2. Methodology

# 1. Efficiency model

. Conditional and robust order-m model (Cazals et al. 2002; Daraio and Simar, 2005).

Order-m approach creates a partial frontier that envelops only m observations from the sample. This procedure is repeated B times resulting in multiples efficiency scores from which the final efficiency measure is computed as the simple mean.

Set of inputs

 $x \in \mathbb{R}^p_+$ 

Set of heterogeneous outputs

 $y \in \mathbb{R}^q_+$ 

Non-discretionary factors

 $z \in \mathbb{R}^r$ 

Production set (without Z)
$$\Psi_m(x) = \left\{ (x',y) \in \mathbb{R}_+^{p+q} \mid x' \leq x, Y_i \leq y, i = 1, ..., m \right\}$$

$$\Psi_m(x) = \left\{ (x',y) \in \mathbb{R}_+^{p+q} \mid x' \leq x, Y_i \leq y, i = 1, ..., m \right\}$$

$$\begin{split} & \Psi_m(x) = \left\{ (x',y) \in \mathbb{R}_+^{p}, \ | \ x \\ & \text{production set (with Z)} \\ & \Psi_m^z(x) = \left\{ (x',y) \in \mathbb{R}_+^{p+q} \mid x' \leq x, Y_1 \leq y, Z = z, i = 1, \dots, m \right\} \end{split}$$



#### · Conditional order-m efficiency model

Conditional model works with probabilistic formulation and incorporates the environmental effect, conditioning the characteristics of the non-discretionary factors.

It constructs a boundary representing the reference set in which each unit is compared. For that, smoothing techniques are needed (Badin et al., 2010; De Witte and Kortelainen, 2013).

$$\mathsf{S'}_{\gamma,n}\left(y\mid x,z\right) = \frac{\sum_{i=1}^{n} I(x_i \leq x,y_i \geq y) K_h(z,z_i)}{\sum_{i=1}^{n} I(x_i \leq x) K_h(z,z_i)}$$

Algorithm (following Daraio and Simar, 2005):

1. Compute the equation (m=100):

$$\tilde{\theta}_{\mathfrak{m}}^{z}\left(x,y\right)=\sup\left\{\theta|\left(x\left|\theta y\right)\right\}\in \Psi_{\mathfrak{m}}^{z}(x)\right\}-E\left[\max_{t=1,\dots,m}\{\min_{j=1,\dots,q}\{\frac{y_{j}^{t}}{y^{d}}\}\right]\mid X\leq x,Z=z\right]$$

2. Redo step 1 for b= 1, ..., B, where B = 200.

3. Finally, 
$$\hat{\theta}_{m,n}(x,y \mid z) \approx \frac{1}{n} \sum_{b=1}^{n} \tilde{\theta}_{m}^{b,z}(x,y)$$
.

#### 2. Spatial model

- · UTM coordinates to validate the geographic location of each school.
- High spatial interdependence among schools.
- Application of SE model (Anselin, 1988):
  - 1. OLS regression:

Betrand = 
$$y + \beta \cdot \hat{\theta}_{ij}^{k,r} + c$$

- 2. Exploratory Spatial Data Analysis (ESDA)
- 3. Tests for spatial dependence detection.
- 4. Fix model step 1 to include spatial dependence detected (ML model).

#### 3. Data and Variables

Data from the Catalan Evaluation Council of the Education System.

n = 1,695 public schools.

t = academic year 2009/2010.



#### 2. Spatial model

a) General regression model

 $y = X\beta + u$   $u \sim N(0, \sigma^2 I)$ 

b) Variants of spatial regression models LAG model

 $y = \rho W y + X \beta + u$ 

 $\varepsilon = \lambda W \varepsilon + u$ 

ERR model

 $y = X\beta + \varepsilon$ Mixed structures

 $y = \rho W_1 y + X \beta_1 + W_2 R \beta_2 + \varepsilon$ 

#### 2. Spatial model

Tests for spatial dependence detection

Spatial dependence tope	Test type	Statistics	Features
	Ad- loc	Moran's I-(Ciff and Ord, 1972) $I = \frac{N}{S} \frac{n'M'n}{n'n}$	c = QLS residues. N = sample size. S : sure of all w <sub>0</sub> W matrix.
Residual	MI.	LM-ERR (Burridge, 1980) $LM = ERR = \frac{\left[e^tWe_{f_{g^2}}\right]^2}{T_1}$	$\mathbf{A}^T\mathbf{m}$ estimation of residual variance. $\mathbf{T}_j\mathbf{m}$ of $(\mathbf{B}^*)\mathbf{W}+(\mathbf{B}^2)$ .
		LM-E1. (Hera and Yoon, 1992) LM = EL $= \frac{e^*We}{[s^2 - T_1(R[_{p-p}]^{-2}e^*Wy/s^2]}^2$ $= \frac{[T_1, T_2^2(R[_{p-p}]^{-2}e^*Wy/s^2]}{[T_1, T_2^2(R[_{p-p}]^{-2}e^*Wy/s^2]}$	$RJ_{\mu\nu\beta} = [T_1 + (WR\overline{\rho})^t M(WR\overline{\rho})/\kappa^2]$ $M = l - K(X^tX)^{-1}N^t$
		$\frac{[\Gamma_3 J_1^{\gamma}(S)_{p=0})^{-1}]}{\text{LM-LAG (Anselin, 1988b)}}$ $LM - LMG = \frac{[a'Wy/s^2]^2}{Rf_{a=0}}$	All the terms are known.
Substantive	ML.	LM-LE (Bern and Yuon, 1992) $LM - LE = \frac{\left[\frac{a^2W_T}{S^2} - a^2Wa/s^2\right]^2}{R_{f_0 = g}^2 - T_L}$	All the terms are known.
iteth	ML	SARMA Test (Anselin, 1988b) $SARMA = \frac{\left[\frac{e'Wy}{s^2} - e'We/s^2\right]^2}{\delta f_{p-q} - T_{q}}$ $= \frac{e'Wy}{s^2 W_{p-q}} = \frac{e'Wy}{s^2 W_{p-q}}$	All the terms are known.
		$+\frac{(\frac{x^2+y^2}{x^2})^2}{T_1}$	

# 1. Efficiency model

 Conditional and robust order-m model (Cazals et al. 2002; Daraio and Simar, 2005).

Order-*m* approach creates a partial frontier that envelops only *m* observations from the sample. This procedure is repeated *B* times resulting in multiples efficiency scores from which the final efficiency measure is computed as the simple mean.

**Set of inputs** 

$$x \in \mathbb{R}^p_+$$

Set of heterogeneous outputs

$$y \in \mathbb{R}^q_+$$

Non-discretionary factors

$$Z \in \mathbb{R}^r$$

Production set (without Z)

$$\Psi_m(x) = \left\{ (x', y) \in \mathbb{R}_+^{p+q} \mid x' \le x, Y_i \le y, i = 1, \dots, m \right\}$$

Production set (with Z)

$$\Psi_{m}^{z}\left(x\right) = \left\{ (x',y) \in \mathbb{R}_{+}^{p+q} \mid x' \leq x, Y_{i} \leq y, Z = z, i = 1, \dots, m \right\}$$

# Conditional order-m efficiency model

Conditional model works with probabilistic formulation and incorporates the environmental effect, conditioning the characteristics of the non-discretionary factors.

It constructs a boundary representing the reference set in which each unit is compared. For that, smoothing techniques are needed (Badin et al., 2010; De Witte and Kortelainen, 2013).

$$S'_{Y,n} \left( y \mid x, z \right) = \frac{\sum_{i=1}^{n} I(x_i \le x, y_i \ge y) K_h(z, z_i)}{\sum_{i=1}^{n} I(x_i \le x) K_h(z, z_i)}$$

Algorithm (following Daraio and Simar, 2005):

1. Compute the equation (m=100):

$$\tilde{\theta}_{m}^{z}\left(x,y\right)=\sup\left\{\theta\left|\left(x\left|\theta y\right)\in\Psi_{m}^{z}(x)\right\}\right|=E\left[\max_{i=1,\dots,m}\left\{\min_{j=1,\dots,q}\left(\frac{Y_{i}^{j}}{y^{j}}\right)\right\}\left|X\leq x,Z=z\right]\right\}$$

2. Redo step 1 for b = 1, ..., B, where B = 200.

3. Finally, 
$$\widehat{\theta}_{m,n}(x,y\mid z) \approx \frac{1}{B} \sum_{b=1}^{B} \widetilde{\theta}_{m}^{b,z}(x,y)$$
.

# 2. Spatial model

- UTM coordinates to validate the geographic location of each school.
- High spatial interdependence among schools.
- Application of SE model (Anselin, 1988):
  - 1. OLS regression:

$$Demand = \gamma + \beta * \tilde{\theta}_m^{b,z} + \varepsilon$$

- 2. Exploratory Spatial Data Analysis (ESDA)
- 3. Tests for spatial dependence detection.
- 4. Fix model step 1 to include spatial dependence detected (ML model).

# 2. Spatial model

# **Tests for spatial dependence detection**

	Spatial dependence type	Test type	Statistics	Features
		Ad- hoc	Moran's I (Cliff and Ord, 1972) $I = \frac{N}{S} \frac{e'We}{e'e}$	$e = \text{OLS residues}$ . $N = \text{sample size}$ . $S = \text{sum of all } w_{ij} W \text{ matrix}$ .
	Residual	ML	LM-ERR (Burridge, 1980) $LM - ERR = \frac{\left[e'We/_{S^2}\right]^2}{T_1}$ LM-EL (Bera and Yoon, 1992)	$s^2$ = estimation of residual variance. $T_1$ = $tr$ $(W'W + W^2)$ .
			LM - EL	$RJ_{\rho-\beta} = [T_1 + (WX\beta)'M(WX\beta)/s^2]$ $M = I - X(X'X)^{-1}X'$
_			$= \frac{\left[\frac{e'We}{s^2 - T_1(RJ_{\rho-\beta})^{-1}e'Wy/s^2}\right]^2}{\left[T_{1-}T_1^2(RJ_{\rho-\beta})^{-1}\right]}$	
			LM-LAG (Anselin, 1988b) $LM - LAG = \frac{[e'Wy/s^2]^2}{RJ_{\rho-\beta}}$	All the terms are known.
_	Substantive	ML	LM-LE (Bera and Yoon, 1992) $LM - LE = \frac{\left[\frac{e'Wy}{s^2} - e'We/s^2\right]^2}{RJ_{\rho-\beta} - T_1}$	All the terms are known.
	Both	ML	SARMA Test (Anselin, 1988b) $SARMA = \frac{\left[\frac{e'Wy}{s^2} - e'We/s^2\right]^2}{RJ_{\rho-\beta} - T_1} + \frac{\left(\frac{e'We}{s^2}\right)^2}{T_1}$	All the terms are known.

# 2. Spatial model



$$y = X\beta + u \qquad u \sim N(0, \sigma^2 I)$$

$$u \sim N(0, \sigma^2 I)$$

### b) Variants of spatial regression models LAG model

$$y = \rho W y + X \beta + u$$

$$W = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1N} \\ w_{21} & 0 & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & 0 \end{bmatrix}$$

**ERR** model

$$y = X\beta + \varepsilon$$

$$\varepsilon = \lambda W \varepsilon + u$$

**Mixed structures** 

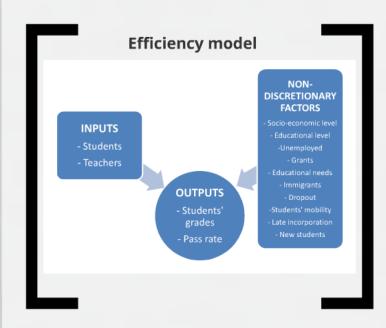
$$y = \rho W_1 y + X \beta_1 + W_2 R \beta_2 + \varepsilon$$

### 3. Data and Variables

Data from the Catalan Evaluation Council of the Education System.

n = 1,695 public schools.

t = academic year 2009/2010.



Sp	atial va	ariables
ı	Variable	Description
$\theta_{cond}$	Conditional Efficiency score	Efficiency index; reflects the school's performance controlling for environmental variables
Area	Territorial area	Specific area within the public education network.
Demand	School demand	Enrollment applications / places offered
Popul	Population	Number of inhabitants
	θ <sub>cond</sub> Area Demand	θ <sub>cond</sub> Efficiency score Area Territorial area Demand School demand

# **Efficiency model**

### **INPUTS**

- Students
- Teachers

## **OUTPUTS**

- Students' grades
- Pass rate

### NON-DISCRETIONARY FACTORS

- Socio-economic level
  - Educational level
    - -Unemployed
      - Grants
  - Educational needs
    - Immigrants
    - Dropout
  - -Students' mobility
  - Late incorporation
    - New students

# **Spatial variables**

Туре	ve Variable		Description		
Independent $ heta_{cond}$		Conditional Efficiency score	Efficiency index; reflects the school's performance controlling for environmental variables		
Area		Territorial area	Specific area within the public education network.		
Dependent	Demand	School demand	Enrollment applications / places offered		
Necessary to separate the sample	Popul	Population	Number of inhabitants		

S residues, N = surples size, S = d ×<sub>g</sub> is trains</sub>, <math>d ×<sub>g</sub> is trains.

and  $v_g$  is trained or foresisted variations.  $T_g$  is P =  $P_g$  +  $(P/2f_g^2)^2M(P/2f_g^2)$ ,  $\phi^2$  =  $-f_g(P/2f_g^2)^2M(P/2f_g^2)$ ,  $\phi^2$  =  $-f_g(P/2f_g^2)^2M(P/2f_g^2)$ .

Let the size interest.</sub>

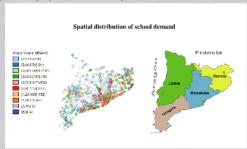
# 3. Empirical results

#### a. Conditional order-m efficiency score

Variable	N	Min	Q <sub>25</sub>	Mean	S.D.	Median	Q <sub>75</sub>	Max
$\theta_{emd}$	1,695	0.98	1.01	1.20	0.01	1.1	1.21	1.25

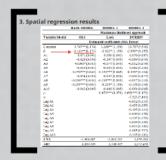
#### b. Spatial study

1. Exploratory Spatial Data Analysis



	ssion res	MODEL 1	Margar 2
		Maximum Bloth	acd approach
Verteile Stodel	00.8	LAG	DUBBLY
	Dete	and estiticion (i	of Cooper
Countral	3,349+43,156	3389**(1.159)	10.7078.334
8	3.110**(1.71)	-2:623*(1.156)	-2,689*(1,158
AL	0.071 0.0415	# E.05630.84E3	8.07590.867
A1	-9 020 X 346	-0.197(0.000)	-9.000 (0.044
A4	-0.000° (C.310)	-6.078 (0.884)	-0.082(0.048
A4	-0.041 (0.041)	-6.038 (D.861)	2.082(0.116
M		-0.310***(D.809)	-9.139*(0.114
AT	-0.055 (0.043)	-0.045 (0.042)	-0.003/00.045
AS .		40.000****(0.045)	-8.017(0.54)
A9	-0.295***(0.061)	-0.225*(1.150)	-8.135(0.548
AD	-0.045 (0.045)	-0.445 (0.845)	-8.054-00.848
P. Contract	$\rightarrow$	E-875 con (0-975)	
			7.575(4.40)
Lag At			-0.012(0.018
Lag A2			2.155(0.355
Light			-0.165(0.54)
Lugas			-0.049(0.15)
Ligas			0.14700.155
Log AC			8.099(0.15)
Log All			3.012(0.182
Log 97			2.025(0.717
Lightin			\$ 11500.040
LNL	-1,090,007	+1,012,223	1,076.00
AXC	2,214,125	2,198,447	2,716.65

	BASE MODEL	MODEL I	MODEL 2
		Madheum Modi	description for the sale
Variable blocket	OLS	Lag	DURBEN
	Estin	nied osefficients (8	M. Erreck
Cerebral	5,707**(1,170)	3.368**(1.350)	10.787(5.536)
	45-1100M (1-171)	(2.6235)1.150)	-2.6591(1.155)
Al	0.871 (0.041)	0.096 (0.840)	0.876 (0.067)
A1	C 313 (3.046)	-9.187 (0.048)	-0.009 (3.044)
A3	-0.000*(0.045)	-9.078 (0.886)	-0.812 (3.049)
61	C 341 (3.041)	-9.034(0.841)	0.002 (0.104)
M	-C 279*** (0.047)	-0.101*** (0.007)	-0.2391(3.114)
AT	C216(0.00)	-0.045 (0.042)	-0.033(0.047)
AS.	4013911-00010	-0.105**** (0.045)	-0.017 (0.148)
40	4 20911 (0.091)	40.2337 [1.150]	-0.139(0.140)
A10	-C211(3,045)	-0.048 (0.049)	-0.811(3.010)
g .		0.073***(0.079)	0.000***(0.070)
			<1.523 (S.481)
LIGAL.			<ul> <li>+0.012 (0.128)</li> </ul>
Log A2			0.455 (0.165)
Lugar		_	-0.195 (0.140)
Lug AS			-0.849 (0.155)
Log AS			0.147 (0.125)
Dog A.7			0.038 (0.150)
Lug Ali			0.812 (0.182)
Lug AS			-0.885 (0.212)
LugA10			0.135 (0.161)
LNL	-0,089,097	-1,012,223	1,079.100
AIC	2.214.135	2,196,447	2,316,696



# 3. Spatial regression results Schools demend index depends on: The efficiency of the school. The area where it is operating. The neighboring school demand index.

#### 1. Exploratory Spatial Data Analysis

Voriable Statistic S.D. Statistic S.D.  Demund 0.54(1)**** 8.91(4) 0.5974*** 7.3894  NOTE WOMAN Marin  **** 0.015 significant fired
NOTE W-Distance Marris
*** = 0.1% significance level

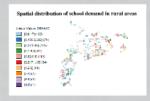
#### 2. Tests for spatial dependence's detection

	Estimated coefficient	p-value
Lagrange multiplier tests		
LM ERR	48.115*	0.04
LMLAG	51.443***	0.000
RLM ERR	0.040	0.841
RLM LAG	3.368**	0.006
SARMA	51.483***	0.000
*= 5% significance level		
** = 1% significance level		
*** = 0.1% significance level		

#### 3. Spatial regression results

Rural vs urban (Eurostat Criterion)
 Rural < 5,000 inhabitants
 Urban >= 5,000 inhabitants

	Ru	urol	Urban		
Tests	Statistic	p-value	Statistic	p-value	
Moran's I	0.46	0.64	0.044*	0.030	
Getis and Ord's G	0.02	0.05	0.091	0.261	
LM ERR	0.66	0.41	0.738	0.342	
LM LAG	0.55	0.41	1.181*	0.031	
RLM ERR	0.15	0.70	0.607	0.560	
RLM LAG	0:04	0.85	0.454*	0.040	
SARMA	0.69	0.71	1.192*	0.044	



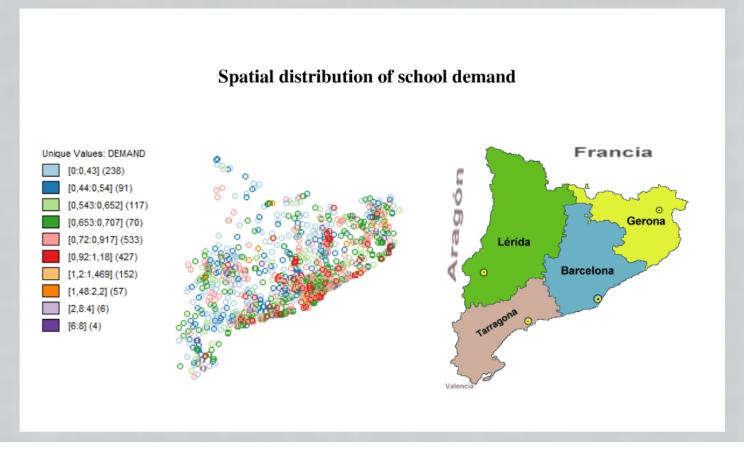
ks for your tention

# a. Conditional order-m efficiency score

Variable	N	Min	$Q_{25}$	Mean	S.D.	Median	$\mathbf{Q}_{75}$	Max
$ heta_{cond}$	1,695	0.98	1.01	1.20	0.01	1.1	1.21	1.25

# b. Spatial study

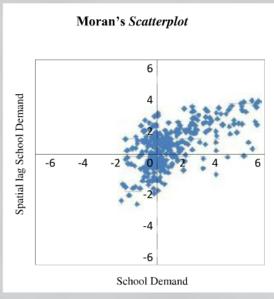
### 1. Exploratory Spatial Data Analysis



# 1. Exploratory Spatial Data Analysis

	Moran	's I	Getis and C	Ord's G		
Variable	Statistic	S.D.	Statistic	S.D.		
Demand	0.3403***	8.9141	0.5974***	7.3894		
NOTE	W=Distance Matrix					

<sup>\*\*\* =</sup> 0.1% significance level



## 2. Tests for spatial dependence's detection

	Estimated coefficient	p-value
Lagrange multiplier tests		
LM ERR	48.115*	0.04
LM LAG	51.443***	0.000
RLM ERR	0.040	0.841
RLM LAG	3.368**	0.006
SARMA	51.483***	0.000

<sup>\*= 5%</sup> significance level \*\* = 1% significance level

<sup>\*\*\* = 0.1%</sup> significance level

	BASE MODEL	MODEL 1	MODEL 2	
		Maximum likelihood approach		
Variable/Model	OLS	LAG	DURBIN	
	Estim	ated coefficients (S	Std. Error)	
Constant	3.747**(1.174)	3.268**(1.159)	10.787 (5.536)	
β	-3.110**(1.171)	-2.622*(1.156)	-2.659*(1.155)	
A1	0.071 (0.041)	0.056 (0.041)	0.076 (0.067)	
A2	-0.020 (0.046)	-0.197 (0.045)	-0.009 (0.046)	
A3	-0.098*(0.045)	-0.074 (0.044)	-0.052 (0.049)	
A5	-0.043 (0.041)	-0.034 (0.041)	0.082 (0.116)	
A6	-0.270***(0.042)	-0.191***(0.045)	-0.239*(0.114)	
A7	-0.056 (0.043)	-0.045 (0.042)	-0.033 (0.047)	
A8	-0.156***(0.044)	-0.105***(0.045)	-0.017 (0.143)	
A9	-0.295***(0.061)	-0.235*(1.156)	-0.138 (0.148)	
A10	-0.043 (0.045)	-0.448 (0.045)	-0.054 (0.048)	
ρ		0.673***(0.070)	0.683***(0.073)	
α			-7.523 (5.401)	
Lag A1			-0.012 (0.135)	
Lag A2			0.455 (0.355)	
Lag A3			-0.195 (0.141)	
Lag A5			-0.049 (0.158)	
Lag A6			0.147 (0.155)	
Lag A7			0.039 (0.150)	
Lag A8			0.012 (0.182)	
Lag A9			-0.085 (0.212)	
Lag A10			0.135 (0.161)	
LNL	-1,096.067	-1,082.223	1,076.330	
AIC	2,214.135	2,188.447	2,316.659	

	BASE MODEL	MODEL 1	MODEL 2	
		Maximum likeli	hood approach	
Variable/Model	OLS	LAG	DURBIN	
	Estimated coefficients (Std. Error)			
Constant	3.747**(1.174)	3.268**(1.159)	10.787 (5.536)	
β	-3.110**(1.171)	-2.622*(1.156)	-2.659*(1.155)	
A1	0.071 (0.041)	0.056 (0.041)	0.076 (0.067)	
A2	<b>-</b> 0.020 (0.046)	-0.197 (0.045)	<b>-</b> 0.009 (0.046)	
A3	-0.098*(0.045)	-0.074 (0.044)	-0.052 (0.049)	
A5	-0.043 (0.041)	-0.034 (0.041)	0.082 (0.116)	
A6	-0.270***(0.042)	-0.191***(0.045)	-0.239*(0.114)	
A7	-0.056 (0.043)	-0.045 (0.042)	-0.033 (0.047)	
A8	-0.156***(0.044)	-0.105***(0.045)	-0.017 (0.143)	
A9	-0.295***(0.061)	-0.235*(1.156)	-0.138 (0.148)	
A10	-0.043 (0.045)	-0.448 (0.045)	-0.054 (0.048)	
ρ		0.673***(0.070)	0.683***(0.073)	
α			-7.523 (5.401)	
Lag A1			-0.012 (0.135)	
Lag A2			0.455 (0.355)	
Lag A3			-0.195 (0.141)	
Lag A5			-0.049 (0.158)	
Lag A6			0.147 (0.155)	
Lag A7			0.039 (0.150)	
Lag A8			0.012 (0.182)	
Lag A9			-0.085 (0.212)	
Lag A10			0.135 (0.161)	
LNL	-1,096.067	-1,082.223	1,076.330	
AIC	2,214.135	2,188.447	2,316.659	

	BASE MODEL	MODEL 1	MODEL 2	
		Maximum likeli	hood approach	
Variable/Model	OLS	LAG	DURBIN	
	Estimated coefficients (Std. Error)			
Constant	3.747**(1.174)	3.268**(1.159)	10.787 (5.536)	
β	-3.110**(1.171)	-2.622*(1.156)	-2.659*(1.155)	
A1	0.071 (0.041)	0.056 (0.041)	0.076(0.067)	
A2	-0.020 (0.046)	-0.197 (0.045)	<b>-</b> 0.009 (0.046)	
A3	<b>-</b> 0.098*(0.045)	-0.074 (0.044)	<b>-</b> 0.052 (0.049)	
A5	-0.043 (0.041)	-0.034 (0.041)	0.082 (0.116)	
A6	-0.270***(0.042)	-0.191***(0.045)	-0.239*(0.114)	
A7	-0.056 (0.043)	-0.045 (0.042)	-0.033 (0.047)	
A8	-0.156***(0.044)	-0.105***(0.045)	-0.017 (0.143)	
A9	-0.295***(0.061)	-0.235*(1.156)	-0.138 (0.148)	
A10	-0.043 (0.045)	-0.448 (0.045)	-0.054 (0.048)	
ρ		0.673***(0.070)	0.683***(0.073)	
α			-7.523 (5.401)	
Lag A1			-0.012 (0.135)	
Lag A2			0.455 (0.355)	
Lag A3			-0.195 (0.141)	
Lag A5			-0.049 (0.158)	
Lag A6			0.147 (0.155)	
Lag A7			0.039 (0.150)	
Lag A8			0.012 (0.182)	
Lag A9			-0.085 (0.212)	
Lag A10			0.135 (0.161)	
LNL	-1,096.067	-1,082.223	1,076.330	
AIC	2,214.135	2,188.447	2,316.659	

	BASE MODEL	MODEL 1	MODEL 2	
		Maximum likelihood approach		
Variable/Model	OLS	LAG	DURBIN	
	Estim	ated coefficients (S	Std. Error)	
Constant	3.747**(1.174)	3.268**(1.159)	10.787 (5.536)	
β	-3.110**(1.171)	-2.622*(1.156)	-2.659*(1.155)	
A1	0.071 (0.041)	0.056 (0.041)	0.076 (0.067)	
A2	-0.020 (0.046)	-0.197 (0.045)	-0.009 (0.046)	
A3	-0.098*(0.045)	-0.074 (0.044)	-0.052 (0.049)	
A5	-0.043 (0.041)	-0.034 (0.041)	0.082 (0.116)	
A6	-0.270***(0.042)	-0.191***(0.045)	-0.239*(0.114)	
A7	-0.056 (0.043)	-0.045 (0.042)	-0.033 (0.047)	
A8	-0.156***(0.044)	-0.105***(0.045)	-0.017 (0.143)	
A9	-0.295***(0.061)	-0.235*(1.156)	-0.138 (0.148)	
A10	-0.043 (0.045)	-0.448 (0.045)	-0.054 (0.048)	
ρ		0.673***(0.070)	0.683***(0.073)	
α			-7.523 (5.401)	
Lag A1			-0.012 (0.135)	
Lag A2			0.455 (0.355)	
Lag A3		_	-0.195 (0.141)	
Lag A5			-0.049 (0.158)	
Lag A6			0.147 (0.155)	
Lag A7			0.039 (0.150)	
Lag A8			0.012 (0.182)	
Lag A9			-0.085 (0.212)	
Lag A10			0.135 (0.161)	
LNL	-1,096.067	-1,082.223	1,076.330	
AIC	2,214.135	2,188.447	2,316.659	

School's demand index depends on:

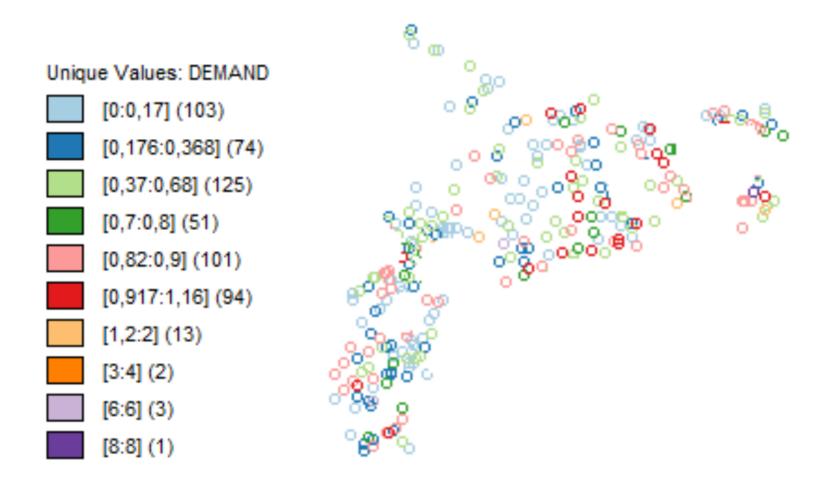
- The efficiency of the school.
- The area where it is operating.
- The neighboring school demand index.

- · Rural vs urban (Eurostat Criterion):
  - Rural < 5,000 inhabitants</li>
  - Urban >= 5,000 inhabitants

	Rural		Urban	
Tests	Statistic	p-value	Statistic	p-value
Moran's I	0.46	0.64	0.044*	0.030
Getis and Ord's G	0.02	0.05	0.091	0.261
LM ERR	0.66	0.41	0.738	0.342
LM LAG	0.55	0.41	1.181*	0.031
RLM ERR	0.15	0.70	0.607	0.560
RLM LAG	0.04	0.85	0.454*	0.040
SARMA	0.69	0.71	1.192*	0.044

<sup>\* = 5%</sup> significance level

# Spatial distribution of school demand in rural areas



# 4. Conclusions

#### Summary

- We use a specific approach different from the previous one in two major aspects:
- First study to offer an analysis of the role of location in explaining the spatial distribution of the school demand in the Spanish context.
- Use of spatial Econometrics techniques combined with non-parametric and robust efficiency estimations.

#### **Implications**

- Provides valuable information to apply improvement programs in less demanded schools.
- Policy makers can introduce additional efforts to reduce the differences among the regions.
  - An active school quality policy might not only affect to one school, but also schools in adjacent zones.

#### Limitations

- Spatial dependence observed may be partially affected by other neighboring private schools that receive public funds.
- · Lack of student level data.
- · Availability of information for several years.

Thanks for attention

# Summary

 We use a specific approach different from the previous one in two major aspects:

 First study to offer an analysis of the role of location in explaining the spatial distribution of the school demand in the Spanish context.

 Use of spatial Econometrics techniques combined with non-parametric and robust efficiency estimations.

# **Implications**

- Provides valuable information to apply improvement programs in less demanded schools.
- Policy makers can introduce additional efforts to reduce the differences among the regions.
  - An active school quality policy might not only affect to one school, but also schools in adjacent zones.

# Limitations

- Spatial dependence observed may be partially affected by other neighboring private schools that receive public funds.
- Lack of student level data.
- Availability of information for several years.

# Thanks for your attention

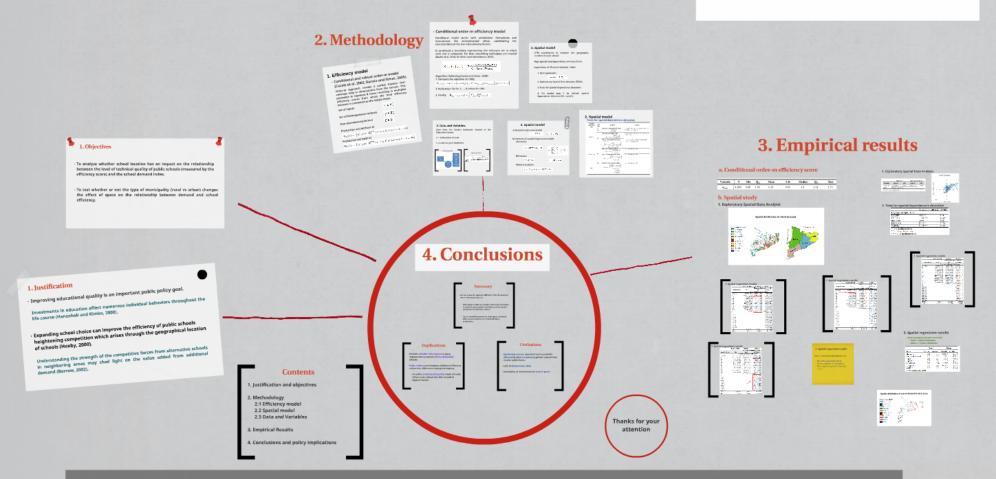












Measuring school demand in the presence of spatial dependence. A conditional approach

Laura López-Torres and Diego Prior