

L'utilisation de données pour analyser les déterminants locaux de la COVID-19

Un exemple à partir de l'article
Covid-19 : analyse spatiale de l'influence des facteurs socio-économiques sur la prévalence et les conséquences de l'épidémie dans les départements français

Par Mounir Amdaoud, Giuseppe Arcuri et [Nadine Levratto](#)



Objectif de la recherche

Analyse des déterminants socio-économiques des hospitalisations, des taux de décès liés à la Covid-19 et de la surmortalité.

Approche territoriale de ces questions grâce au recours à des données calculées au niveau des départements français.

Champ de la géographie de la santé qui a émergé dans les années 1970, dans le but de rendre compte des disparités spatiales en matière d'accès aux soins, d'exposition aux maladies et de morbidité (Rosenberg, 1998 ; Curtis 2004 ; Prior et I., 2018).

Une analyse spatiale exploratoire qui révèle l'hétérogénéité et l'autocorrélation spatiale de la maladie et de ses conséquences.

Recours à des modèles d'économétrie spatiale permettant de mettre en évidence l'influence des caractéristiques territoriales

Introduction (1)

L'apparition de l'épidémie en France et la dimension régionale

L'importance des facteurs socioéconomiques locaux comme facteur explicatif de l'état de santé des populations et des taux de mortalité est démontrée par la littérature (Cambois et Jusot, 2007)

La prise en considération d'éléments de contexte dans l'analyse d'une question liée à la santé est nécessaire à plusieurs niveaux:

- les différences d'un territoire à l'autre en matière de santé ou de maladie sont de plus en plus souvent pointées par la littérature
- la responsabilité des différents types de caractéristiques d'un territoire peuvent jouer différemment selon le type d'affection
- l'organisation du système de santé est régionalisée

Introduction (2)

Les termes du débat étudiés dans la nouvelle version de l'article:

- La situation des hôpitaux publics et de la médecine de ville
- Le rôle du respect des institutions (van Bavel et al., 2020) et la confiance sociale (Barrios et al., 2021) soulignés pour expliquer le respect des règles de distanciation et le confinement.

Démarche empirique reposant sur les cartes de la diffusion de la Covid-19 par département élaborées à partir de communiquées par l'INSEE et le Ministère des Solidarités et de la Santé depuis le début du mois de Mars 2020

Mobilisation de différentes variables locales pouvant potentiellement affecter le taux de malades et le taux de mortalité liés à la Covid-19

van Bavel J.J., Baicker K., Boggio P.S., et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav* 2020; 4: 460–471.

<https://doi.org/10.1038/s41562-020-0884-z>

Barrios J. M., Benmelech E., Hochberg Y.V., Sapienza P., Zingales L. Civic capital and social distancing during the Covid-19 pandemic. *J Public Econ* 2021; 193.

<https://doi.org/10.1016/j.jpubeco.2020.104310>

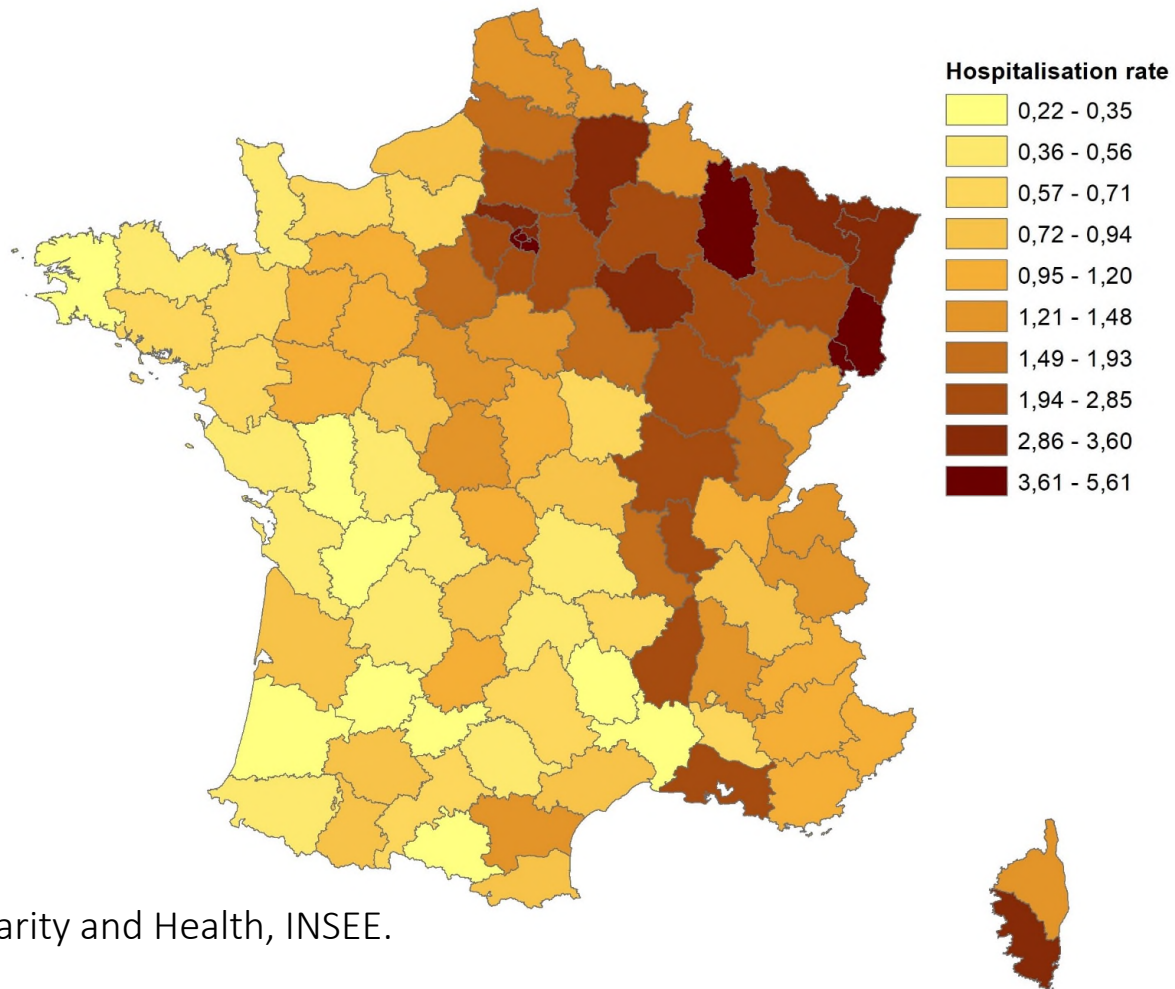
Les indicateurs d'intensité de l'épidémie

Niveau d'analyse: 96 départements de France métropolitaine

3 indicateurs d'intensité de l'épidémie :

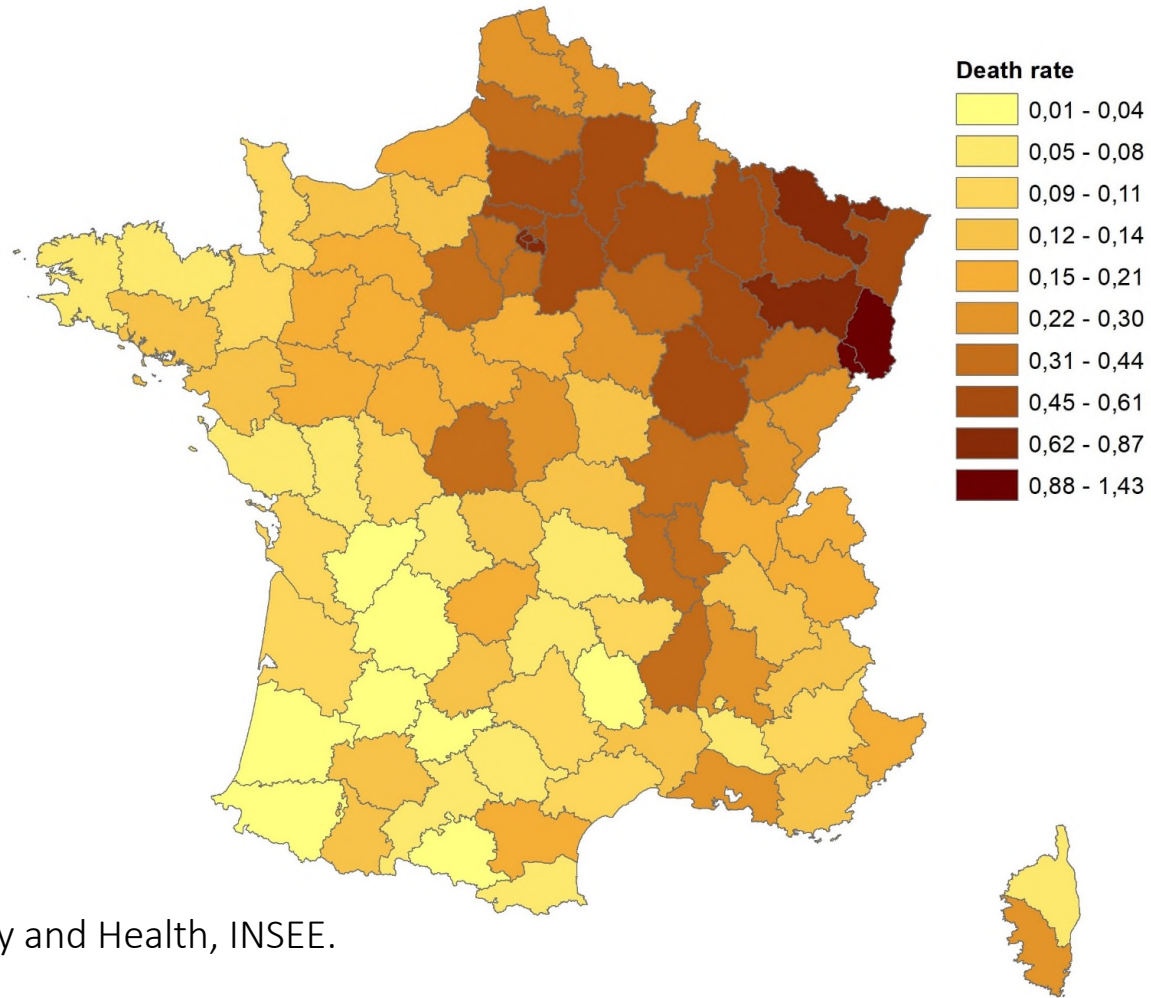
- Taux d'hospitalisation: nouvelles hospitalisations Covid-19 (entre le 19 mars et le 31 août) / n habitants *1000 [*Source: Santé Publique France*];
- le taux de mortalité: nombre cumulé de décès Covid-19 à l'hôpital (entre le 19 mars et le 31 août)/ n habitants*1000 [*Source: Santé Publique France*];
- un taux de surmortalité : différence entre le nombre total de décès par département (entre le 1er mars et le 24 août) et la moyenne des décès recensés à la même date en 2018 et 2019 rapportée à cette dernière [*Source: INSEE*];

Taux d'hospitalisation Covid-19



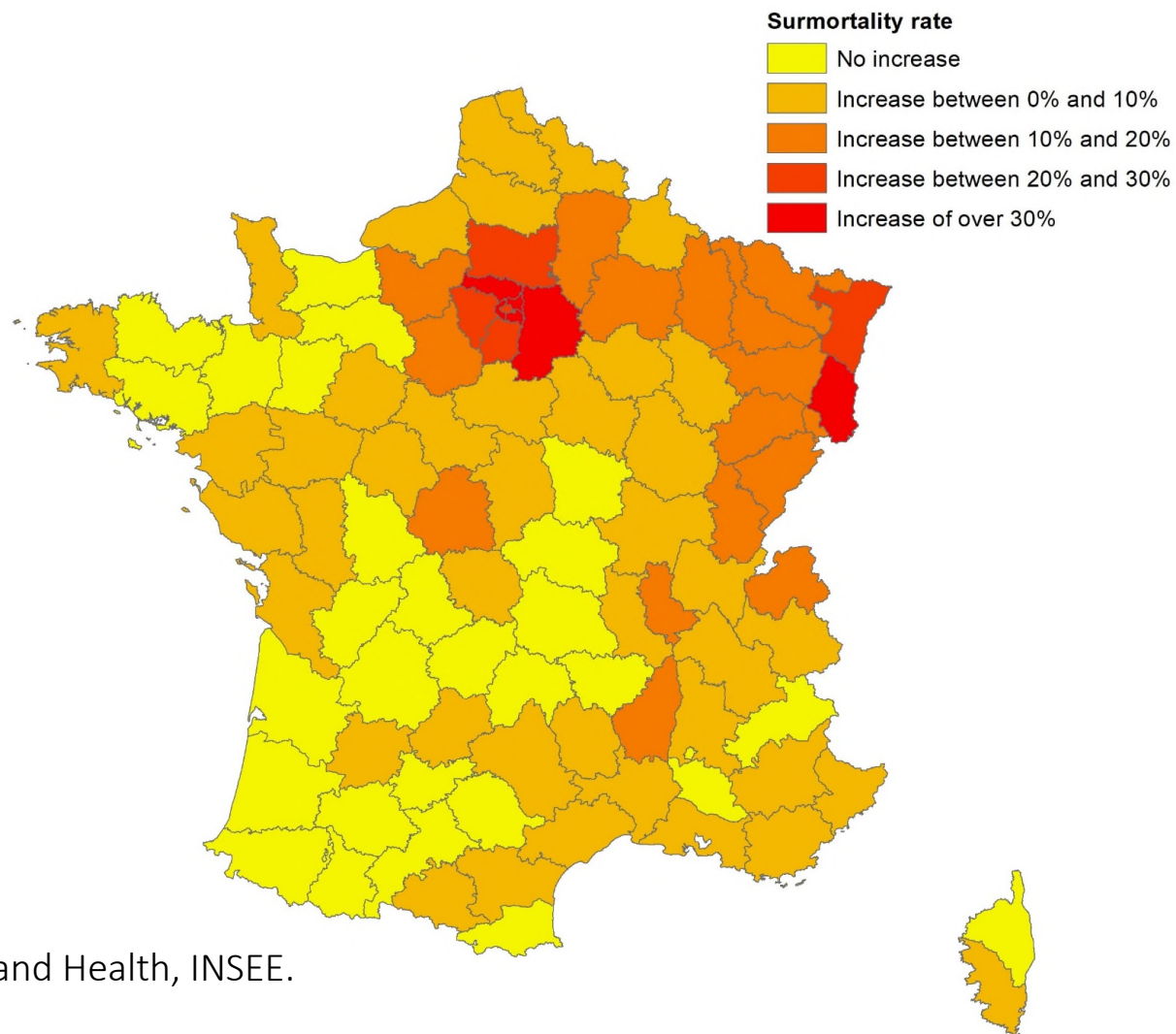
Sources: Ministry of Solidarity and Health, INSEE.
Authors' elaboration.

Taux de décès Covid-19



Sources: Ministry of Solidarity and Health, INSEE.
Authors' elaboration.

Taux de surmortalité



Sources: Ministry of Solidarity and Health, INSEE.
Authors' elaboration.

Analyse exploratoire spatiale (ESDA)

Matrice spatiale W utilisée : celle dite de distance inverse.

L'autocorrélation spatiale est mesurée par la **statistique I de Moran**: valeurs entre -1 et 1:

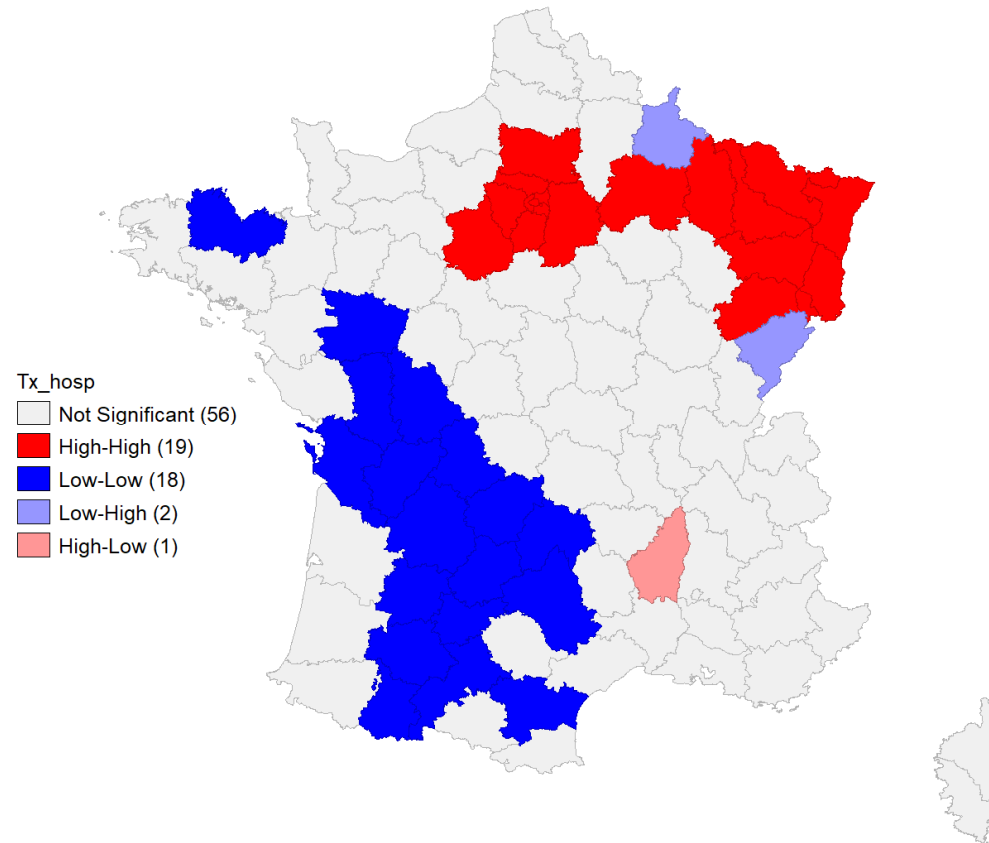
- Valeurs négatives: autocorrélation spatiale négative
- 0: modèle spatial aléatoire
- Valeurs positives: autocorrélation spatiale positive

Indicator	Moran's I	Mean	SD	Standardized value	P-value
Hospitalisation rate on 31-08	0.629	-0.0105	0.0755	8.4684	0.0001
Death rate on 31-08	0.661	-0.0105	0.0750	8.9597	0.0001
Surmortality rate on 24-08	0.737	-0.0105	0.0758	9.8645	0.0001

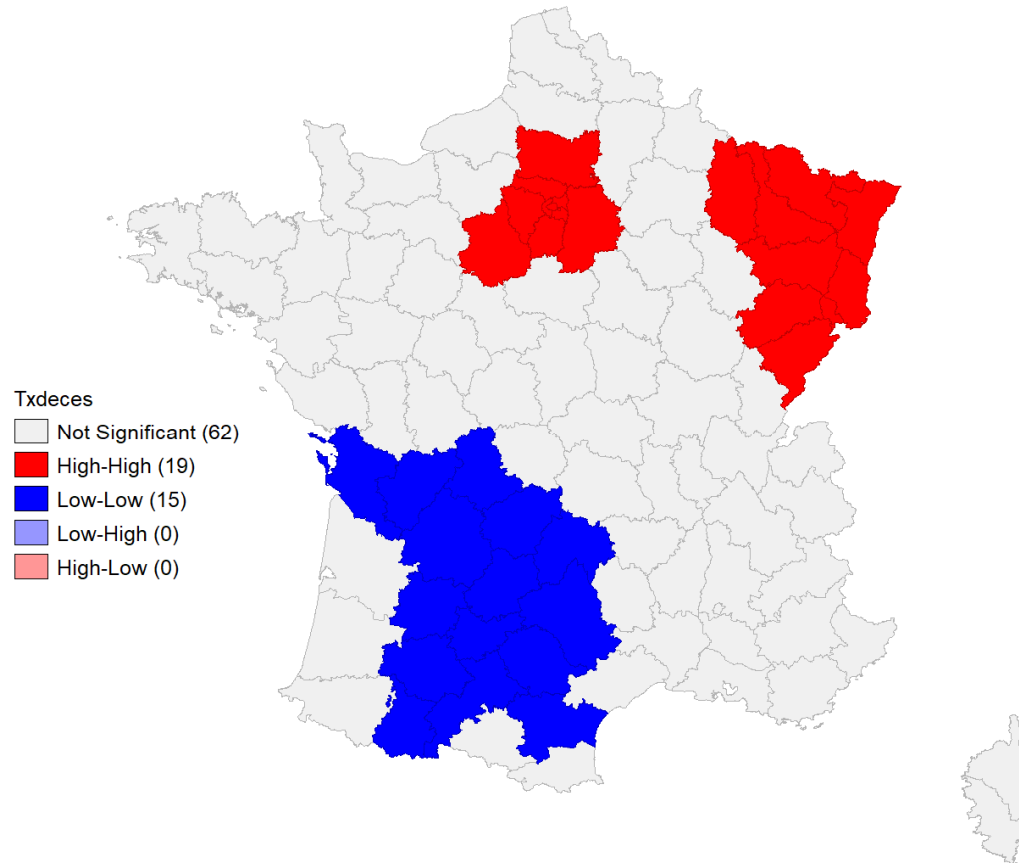
Source: Santé Publique France and INSEE.

Authors' elaboration

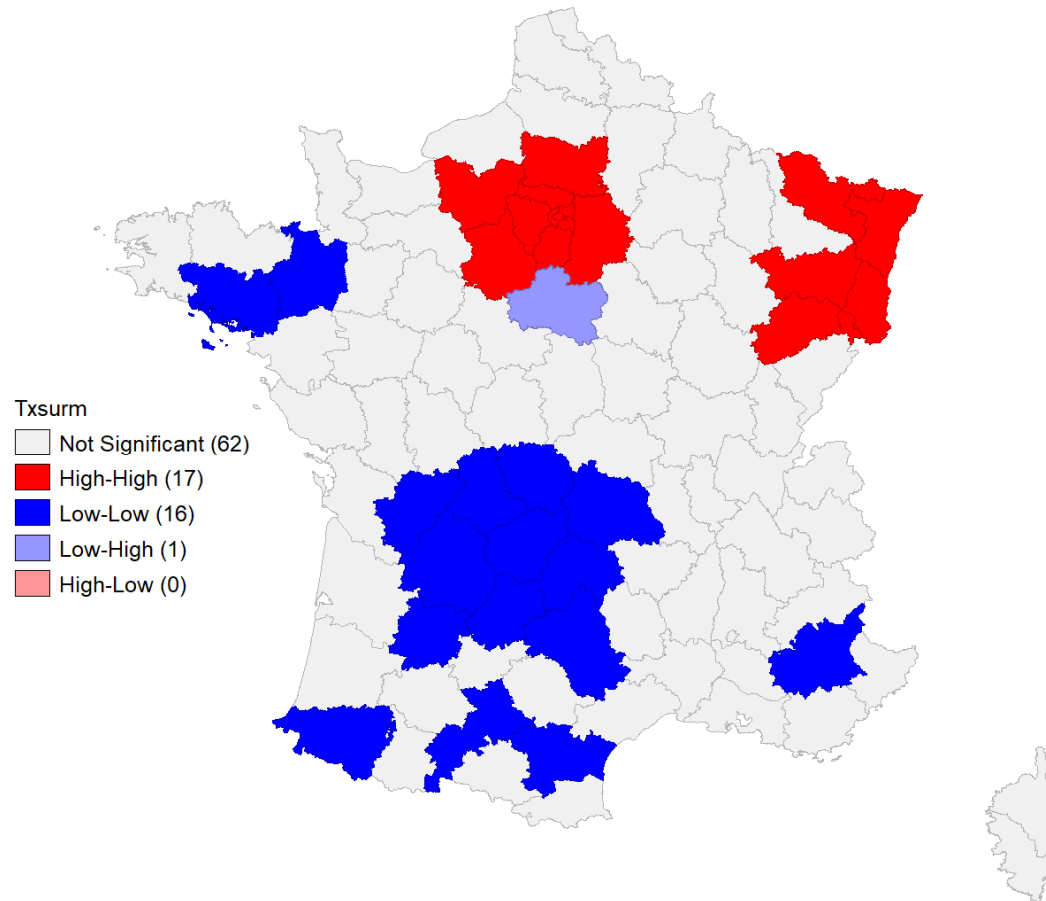
LISA: taux d'hospitalisation



LISA: taux de décès



LISA: taux de surmortalité



Variables explicatives

Explanatory variables			
Index of medical density	Index of medical density	2018	INSEE
Social capital	Participation rate in the 2017 presidential elections in the 1st round	2017	INSEE
Population density	Total population per km ² (log)	2016	INSEE
Concentration rate	Share of the 3 main municipalities in the population of the department	2016	INSEE
Share of the population aged 65 and over	Number of individuals aged 65 and older over total population. The variable equals one if the value is greater or equal to the mean	2016	INSEE
Rate of blue-collars	Number of blue collar workers over total population	2016	INSEE
Median standard of living	Median standard of living in 2016 (log)	2016	INSEE
Rural territory	Dummy variable equals one if the territory is predominantly rural	2020	OCDE

Modèles économétriques

$$\text{MCO} \quad Y = X\beta + \varepsilon$$

Lorsqu'un phénomène d'autocorrélation spatiale est ignoré dans la spécification du modèle, mais présent dans le processus générateur des données, les estimateurs des MCO sont biaisés et non convergents.



$$\text{SAR} \quad Y = \rho WY + X\beta + \varepsilon$$

WY est la variable endogène décalée pour la matrice de distance inverse W , ρ est le paramètre autorégressif indiquant l'intensité de l'interaction existante entre les observations de Y

$$\text{SEM} \quad Y = X\beta + \varepsilon \text{ avec } \varepsilon = \lambda W\varepsilon + u$$

Le paramètre λ reflète l'intensité de l'interdépendance entre les résidus de la régression et u est le terme d'erreur

Modèles économétriques

$$\text{MCO} \quad Y = X\beta + \varepsilon$$

Lorsqu'un phénomène d'autocorrélation spatiale est ignoré dans la spécification du modèle, mais présent dans le processus générateur des données, les estimateurs des MCO sont biaisés et non convergents.

$$\text{SAR} \quad Y = \rho WY + X\beta + \varepsilon$$

Résultats du Lagrange Multiplier test

WY est la variable endogène décalée pour la matrice de distance inverse W , ρ est le paramètre autorégressif indiquant l'intensité de l'interaction existante entre les observations de Y

$$\text{SEM} \quad Y = X\beta + \varepsilon \quad \text{avec} \quad \varepsilon = \lambda W\varepsilon + u$$

Le paramètre λ reflète l'intensité de l'interdépendance entre les résidus de la régression et u est le terme d'erreur

Résultats

Variable	Hospitalisation rate 31-08	Death rate 31-08	Excess mortality rate 24-08
Index of medical density	-1.240**	-0.299**	-0.124***
	(0.623)	(0.123)	(0.0474)
Social capital	-0.115***	-0.0151**	-0.0110***
	(0.0377)	(0.00705)	(0.00302)
Population density	0.337***	0.0804***	0.0362***
	(0.111)	(0.0220)	(0.00861)
Concentration rate	1.935**	0.498***	-0.0488
	(0.845)	(0.166)	(0.0642)
Share of the population aged 65 and over	-0.0641	0.0240	0.00274
	(0.209)	(0.0410)	(0.0158)
Rate of blue-collars	6.889	2.077**	0.543
	(4.667)	(0.929)	(0.352)
Median standard of living	1.881	0.333	0.185*
	(1.292)	(0.255)	(0.0997)
Rural territory	0.0459	-0.0166	1.30e-05
	(0.217)	(0.0427)	(0.0165)
Constant	-10.95	-2.625	-1.090
	(12.37)	(2.453)	(0.943)
ρ	0.470***	0.567***	0.452***
	(0.0883)	(0.0796)	(0.0820)

Conclusions

Au-delà de l'importance des caractéristiques individuelles comme facteurs explicatifs de la probabilité de contracter la Covid-19 et de ses conséquences, les éléments liés au contexte économique, démographique et social interviennent également

- La culture locale est un déterminant significatif de l'épidémie de Covid-19
- Même résultat pour les politiques de santé publique, financées sur fonds publics
- Les départements les plus denses, les plus inégalitaires ainsi que ceux dans lesquels la part d'ouvriers est la plus élevée sont les plus vulnérables
- Les disparités locales et les effets de débordement vont dans le sens d'une régionalisation poussée de la mise en œuvre des politiques de santé

Limites: Nécessité de conduire d'autres recherches sur les causes et conséquences socio-économiques de cette épidémie en raison du caractère évolutif des données mobilisées et des possibles effets retard de la Covid-19

Pour aller plus loin

- Amdaoud M., G. Arcuri et N. Levratto (2021) Are regions equal in adversity ? A spatial analysis of the spread and dynamics of COVID-19 in Europe, *European Journal of Health Economics*, à paraître
- Geography of COVID-19 outbreak and first policy answers in European regions and cities, Rapport pour le programme Européen ESPON, décembre 2020, Coordinateur : S. Bourdin (EM Normandie)