

# Air Pollution and Children's Health Inequalities

Milena Suarez Castillo (DREES, PSE)  
Vianney Costemalle (DREES)  
Christine Le Thi (DREES)  
David Benatia (HEC Montréal, CREST)

Juin 2023  
FIME

## Comment les inégalités d'exposition se traduisent-elles par des inégalités de santé ?

- La mauvaise qualité de l'air donne lieu à des risques de santé, en particulier chez les jeunes enfants.
- Les inégalités d'exposition à la pollution de l'air contribuent aux inégalités de santé dès la naissance.
- Les effets de cette exposition sont hétérogènes à travers les individus, à cause de plusieurs facteurs.
- Pourtant, difficile d'aller plus loin que l'évaluation d'un effet moyen de la pollution sur la santé d'une population.
- Dans cet article, nous explorons l'hétérogénéité des effets et cherchons à identifier ces facteurs, notamment socio-économiques mais aussi liés à la vulnérabilité initiale.

## Problèmes respiratoires chez le jeune enfant: quelques chiffres

- En 2017, les médicaments contre les maladies obstructives des voies respiratoires représentaient 1 milliard d'euros remboursés par l'assurance maladie pour 8.6 millions de bénéficiaires, dont 26.6% de moins de 20 ans.
- Chez les moins de 2 ans, 1/2 enfant a au moins une prescription de médicament contre les maladies respiratoires dans l'année.
- Dans l'échantillon d'étude, en moyenne de 2009 à 2015, un tiers des admissions en urgence des enfants de moins de 3 ans sont des admissions au motif principal liés à des problèmes respiratoires, dont 65% pour bronchiolite et/ou asthme.
- Une proportion significative du taux d'incidence de l'asthme chez l'enfant en Europe attribuable à la pollution atmosphérique, et évitable (Khreis, 2019).

## Qui sont les plus exposés?

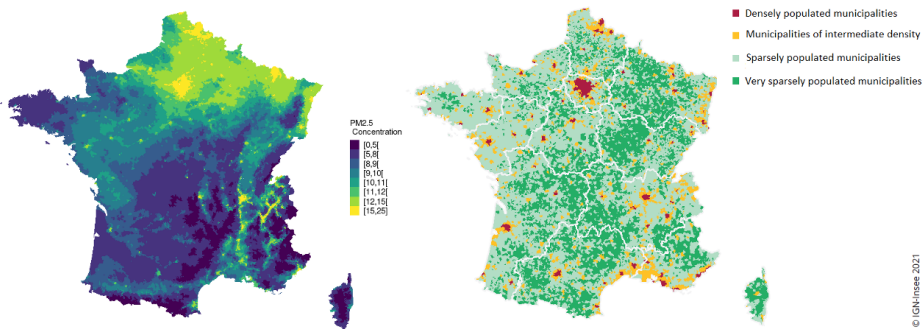


Figure: PM2.5 concentrations ( $\mu\text{g}/\text{m}^{-3}$ ) in 2010 (source: ACAG)

## Qui sont les plus exposés?

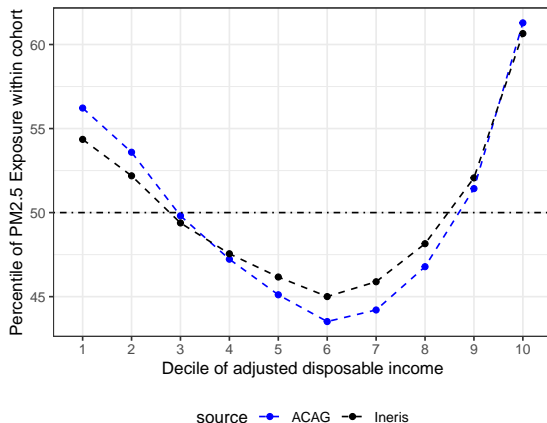


Figure: Exposure on birth year by parental income in France (2016)

# La vulnérabilité au coeur des effets...

## .. mais rarement quantifiée

TABLE 7—SUMMARY STATISTICS FOR THE MEDICARE BENEFICIARIES MOST AND LEAST AFFECTED BY POLLUTION

Outcome:	Bottom 75% (1)	Top 1% (2)	Difference (3)
Life expectancy (years)	11.6	3.65	-7.91 (0.00192)
<i>Demographics</i>			
Age (years)	75.7	82.9	7.20 (0.00321)
<i>Chronic conditions</i>			
Alzheimer's or dementia	0.0849	0.470	0.385 (0.000155)
Chronic kidney disease	0.113	0.422	0.309 (0.000164)
COPD	0.169	0.495	0.326 (0.000197)
Heart failure	0.187	0.650	0.463 (0.000128)
Lung cancer	0.00781	0.0705	0.0627 (0.0000601)
<i>Medical spending (dollars)</i>			
Durable medical equipment	136	785	648 (1.05)

Figure: Most and Least Affected by Pollution [Deryugina et al., 2019]

- Aux U.S., les personnes âgées susceptibles de mourir du fait de la pollution de l'air ne sont pas nécessairement situées dans les endroits les plus pollués (Deryugina et al. 2021)
- Mais elles cumulent les facteurs de risques, et ont des espérances de vie attendues faibles (Deryugina et al. 2019)
- Ces différences sont-elles vraies aussi pour les jeunes enfants?
- Quel est le rapport avec le revenu des parents?

## Estimation directe de cette hétérogénéité?

- **Effet de traitement moyen (ATE)**
- Être vulnérable à la pollution de l'air (ici): aller à l'hôpital/consommer des soins du fait de la pollution de l'air: Idée de causalité fondamentale
- Comparer un groupe d'enfants "traités" par un surcroît (exogène) d'exposition à la pollution de l'air à un groupe d'enfants "contrôles"
- Traitement binaire quasi-expérimental
- E.g. Jans et al. (2018)
- **Effet de traitement moyen conditionnel (CATE)**
- Décrire les groupes d'impacts i.e. les X% les plus affectés par le traitement.
- "Qui sont les jeunes enfants dont la santé respiratoire est la plus affectée par la pollution de l'air?": ce sont les jeunes enfants les plus affectés par le traitement, et on peut les décrire
- Est-ce que le premier décile et le dixième décile y sont sur-représentés?
- E.g. Deyugina et al. (2019), Chernozhukov et al. (2018, 2022)

## Dans cet article

- Échantillon de 336 169 enfants nés en France de 2008 à 2017.
- Caractériser ces enfants en fonction d'un ensemble varié de variables provenant de données administratives.
- Stratégie empirique :
  - ① Définir un **traitement binaire**, crédiblement exogène, augmentant l'exposition à la pollution de l'air pour certains enfants.
  - ② Estimer les **effets moyens du traitement sur les résultats de consommation de santé**.
  - ③ Utiliser l'apprentissage automatique générique [Chernozhukov et al., 2022] pour caractériser ceux qui sont les plus affectés.
- Nos résultats suggèrent que les impacts sont concentrés sur 10% des nourrissons, caractérisés principalement par de mauvais indicateurs de santé mais aussi par un revenu parental plus faible.



## Littérature

- Les inégalités sociales résultant des risques sanitaires liés à la qualité de l'air ambiant [Deguen and Zmirou-Navier, 2010]
- Des preuves quasi-expérimentales sur les risques environnementaux [Dominici et al., 2014, Bind, 2019]
- **Revenu des parents et pollution de l'air** [Jans et al., 2018]
- **Conséquences à long terme de l'exposition en bas âge** [Currie, 2011, Isen et al., 2017]
- **Hétérogénéité par GenericML**, ou inférence agnostique.  
*Théorie* : [Chernozhukov et al., 2018, Chernozhukov et al., 2022, Imai and Li, 2022]. *Application* : [Deryugina et al., 2019, Breda et al., 2018]
- **Estimations causales de la pollution de l'air à l'aide d'instruments liés aux conditions météorologiques** [Arceo et al., 2016, Schwartz et al., 2016, Jans et al., 2018, Deryugina et al., 2019, Dechezleprêtre et al., 2019]

## L'EDP-Santé 2017

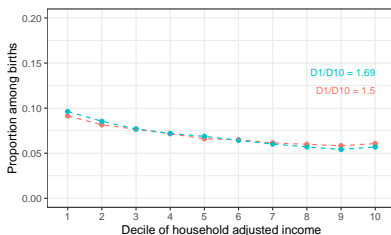
Pour un échantillon de 4% de la population résidente française, les individus ED, caractérisés par leur date de naissance, plusieurs sources de données de référence partagent un identifiant commun permettant leur mise en correspondance :

- La consommation de soins de santé remboursée (*SNDS : PMSI et SNIIRAM*) de 2008 à 2018
- Les informations de résidence au niveau de la commune (*Recensement, données fiscales, registres d'état civil*)
- Les données fiscales des membres du foyer de ces individus (*contrepartie exhaustive : Fidéli, Filosofi*), pour les revenus de 2010 à 2016

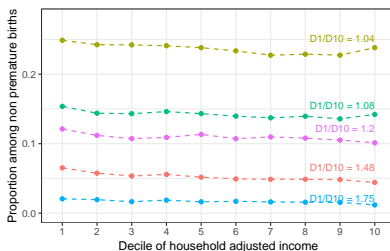
Nous sélectionnons les enfants nés entre 2008 et 2017 ayant au moins une information sur leur résidence et qui sont bénéficiaires de l'Assurance Maladie Nationale, soit 336 169 enfants.

# Caractérisation de l'état de santé de référence avec le séjour à la maternité

La quasi-exhaustivité des naissances en France est enregistrée dans le PMSI (activités hospitalières).



- Premature birth
- Low birth weight (<2.5kg)



- Neonathology at birth
- Birth with a significant issue (GHM)
- Birth with at least one pathological diagnosis (DP des UM)
- Respiratory radiographie at birth
- Prolonged hospital stay at birth (greater than 5 days)

## Variables de santé

À partir du PMSI (activités hospitalières)

- Admissions aux urgences pour l'asthme (respectivement la bronchiolite) avant le 3e anniversaire (respectivement le 2e anniversaire)

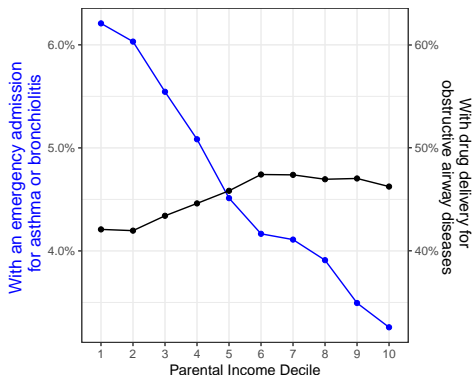
1,2% des nourrissons de l'échantillon ont eu au moins une admission aux urgences pour l'asthme avant leur 3e anniversaire, 3,5% pour la bronchiolite avant leur 2e anniversaire.

À partir du SNIIRAM (livraison de médicaments remboursés)

- Médicaments pour les maladies des voies respiratoires obstructives

26% ont reçu au moins une livraison de ce type de médicaments avant leur 1er anniversaire, 43% avant leur 3e anniversaire.

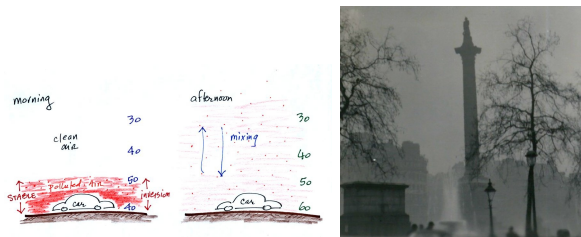
## Variables de santé



**Figure:** Indicateurs de santé respiratoire d'ici le troisième anniversaire par décile du revenu ajusté du ménage

## Les inversions thermiques comme variations exogènes de la pollution de l'air

[Jans et al., 2018, Arceo et al., 2016, Dechezleprêtre et al., 2019]

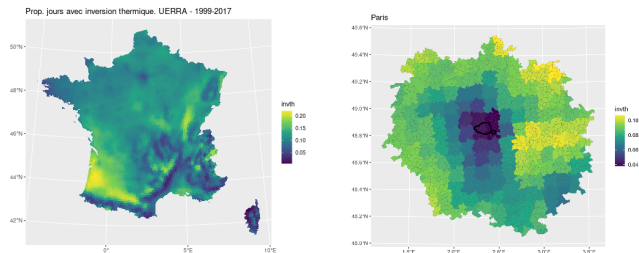


**Figure:** Les inversions thermiques. Gauche : [www.atmo.arizona.edu](http://www.atmo.arizona.edu), Droite : Le grand smog de Londres, 1952, causé par une inversion thermique durable et de fortes émissions de polluants.

Dans cet article : comparer les enfants exposés au cours de leur 1ere année à un nombre anormalement élevé d'inversions thermiques par rapport aux autres enfants.

# Inversions thermiques

Réanalyse régionale UERRA pour l'Europe, produit du programme Copernicus : 11 km x 11 km sur 11 niveaux d'altitude.



Un jour avec une inversion thermique est défini lorsque

$$\langle T \rangle_t(500\text{m}) - \langle T \rangle_t(15\text{m}) > 0$$

## Pollution de l'air : PM2.5

Deux sources de référence [Hammer et al., 2020][Real et al., 2022]

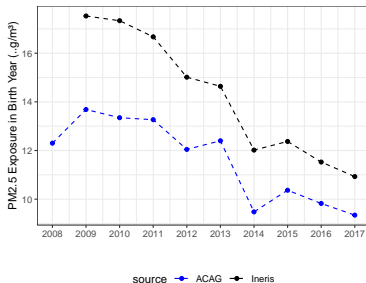
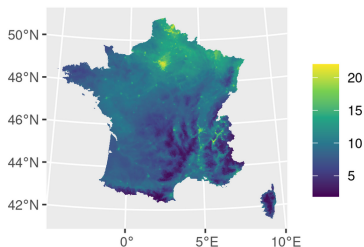


Figure: Concentrations de PM2.5 ( $\mu\text{g}/\text{m}^{-3}$ ) en 2010 (source : ACAG)

Figure: Exposition lors de l'année de naissance selon le revenu des parents en France (2016)

- Sur la période 2008-2017, sur l'ensemble du territoire métropolitain, uniquement disponible par année civile
- Disparités entre les sources
- Les inversions thermiques n'influencent pas spécifiquement ce polluant [Godzinski and Suarez Castillo, 2021], nous interprétons donc nos estimations comme l'impact d'un effet agrégé de la pollution de l'air.



## Un choc quasi-expérimental?

- Pourquoi s'appuyer sur une variation exogène?  $T_i$  est un choc binaire **quasi-expérimental** d'exposition pour résoudre l'endogénéité de la relation entre pollution et santé.
  - ① La pollution atmosphérique est un produit des activités économiques (transports, chauffage, industrie), qui sont corrélées avec/ou affectent la santé des populations. Par exemple, relation pro-cyclique entre tabagisme et condition macro-économique
  - ② Les choix de localisations résidentiels peuvent être influencés par les niveaux de pollution ou par des aménités corrélées (e.g. axe routier majeur). Par exemple, les mobilités résidentielles concomitantes à la naissance d'un enfant sont orientés vers des lieux moins pollués, et ce d'autant plus que le niveau de vie est élevé.
  - ③ Erreur de mesure (e.g. ACAG vs Ineris)

## Définition du groupe traité et non traité

Un nourrisson  $i$  est considéré comme "traité" lorsque le nombre de jours avec inversion thermique locale au cours de sa première année dépasse la moyenne locale à long terme d'une certaine valeur seuil  $\bar{n}$ .

$$T_i^{\bar{n}} = 1 \left\{ N_i - \hat{N}_i \geq \bar{n} \right\}, \quad (1)$$

Un modèle auxiliaire est utilisé pour prédire le nombre attendu de jours avec inversion thermique localement par :

$$N_{ct} = n_c + \delta_t + \epsilon_{ct}$$

## Caractéristiques d'un tel traitement

- ① Basé sur les inversions thermiques : il est crédible que la quantité d'inversions thermiques locales soit sans lien avec la santé des enfants, une fois pris en compte les autres conditions météorologiques en surface.
- ② La probabilité d'être traité est mieux répartie sur le territoire  
Figures
- ③ Les groupes traités et non traités sont similaires, à la fois de manière inconditionnelle et conditionnelle. Tableaux
- ④ Capacité à capturer la surexposition

$$PM_i = \theta_{\bar{n}} T_i^{\bar{n}} + X_i^W \lambda + \rho_c + \xi_t + e_i, \quad (2)$$

où  $PM_i$  mesure les concentrations de particules fines pour l'enfant  $i$  en fonction de sa commune de résidence  $c$  et de son année de naissance  $t$ , en faisant varier  $\bar{n}$ .

## PM2.5 et un nombre anormalement élevé de jours avec inversion thermique.

$$PM_i = \theta_{\bar{n}} \times 1 \left\{ N_i - \hat{N}_i \geq \bar{n} \right\} + X_i^W \lambda + \rho_c + \xi_t + e_i$$

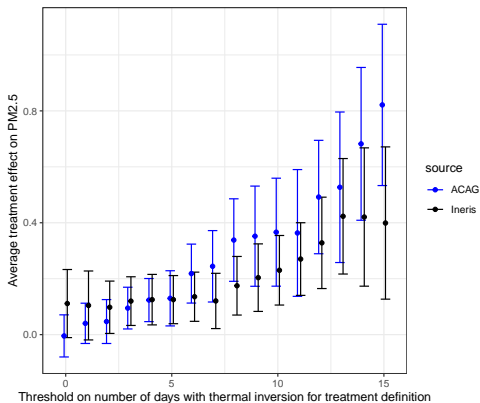


Figure:  $\bar{n} \rightarrow \hat{\theta}_{\bar{n}}$ . Note: Infants born in January.

# Traitement binaire d'exposition

More results

	Intermediate shock of exposure $\bar{n} = 7$	Large shock of exposure $\bar{n} = 10$
<i>Additional days with thermal inversion</i>		
Sample size	10.79 [10.48;11.11] 336169	11.38 [10.92;11.83] 336169
<i>Additional exposure to PM2.5 in first year, by source, for those born in January</i>		
ACAG (2008-2017)		
Sample size	0.12 [0.022;0.22] 82164	0.23 [0.11;0.35] 82164
Ineris (2009-2017)		
Sample size	0.24 [0.12;0.37] 73354	0.37 [0.17;0.56] 73354

## Effets de traitement moyen

$$Y_i = \beta \times 1 \left\{ N_i - \hat{N}_i \geq \bar{n} \right\} + X_i^I \gamma_I + X_i^W \gamma_W + \mu_c + \delta_t + \epsilon_i, \quad (3)$$

- $T_i$ : choc binaire quasi-expérimental d'exposition conditionné par des contrôles, et des effets fixes commune et année.
- $X_i^I$ : caractéristiques de l'enfant<sup>1</sup>.
- $X_i^W$  conditions météorologiques<sup>2</sup>.
- *Hypothèse : être exposé à un nombre anormalement élevé d'inversions thermiques au cours de sa première année, conditionnellement aux contrôles répertoriés, n'est pas lié à la santé, à l'exception de l'impact qu'il a sur la pollution de l'air.*

---

<sup>1</sup> sexe, revenu des parents (exprimé en déciles et de manière linéaire), caractéristiques de la mère (âge et indicateur de naissance à l'étranger), âge gestationnel (indicateur de naissance prématurée et de manière linéaire), poids à la naissance (indicateur de faible poids à la naissance et de manière linéaire), et six autres indicateurs de santé dérivés du séjour hospitalier à la naissance

<sup>2</sup> le nombre de jours dans chacun des 20 intervalles de température observés, le nombre de jours dans chacun des 12 intervalles d'intensité du vent observés, ainsi que des polynômes du deuxième degré pour la pression atmosphérique moyenne et l'humidité

## Effets de traitement moyen

Figures by threshold

$X^I$	Intermediate shock of exposure		Large shock of exposure	
	no	$\bar{n} = 7$ yes	no	$\bar{n} = 10$ yes
<i>Drug delivery for obstructive airway diseases</i>				
- before first anniversary				
	0.003	0.003	0.012	0.012
	[-0.003;0.009]	[-0.004;0.010]	[0.004;0.019]	[0.003;0.020]
Sample size	336169	217859	336169	217859
- between first and third anniversary				
	0.059	0.067	0.085	0.089
	[0.048;0.071]	[0.054;0.079]	[0.068;0.10]	[0.072;0.11]
Sample size	289843	204549	289843	204549
<i>Hospital emergency admissions</i>				
- asthma before third anniversary				
	0.002	0.003	0.002	0.003
	[0.000;0.005]	[0.001;0.006]	[-0.000;0.005]	[-0.000;0.006]
Sample size	246075	189722	246075	189722
- bronchiolitis before second anniversary				
	0.005	0.003	0.005	0.001
	[0.002;0.008]	[-0.000;0.006]	[0.002;0.008]	[-0.002;0.005]
Sample size	278302	207825	278302	207825
- bronchiolitis or asthma before third anniversary				
	0.005	0.007	0.005	0.007
	[0.002;0.009]	[0.003;0.011]	[0.001;0.010]	[0.002;0.013]
Sample size	246075	189722	246075	189722

# Hétérogénéité des effets de traitement

[Chernozhukov et al., 2018]

- CATE - Conditional average treatment effect  
 $s_0(Z) = \mathbb{E}[Y(1) - Y(0)|Z]$
- Absence d'hétérogénéité :  $s_0(Z) = s_0 = \mathbb{E}[Y(1) - Y(0)]$
- Soit  $S(Z)$  un proxy de  $s_0(Z)$ , appris des données ("Generic")

[Chernozhukov et al., 2018] propose une méthode d'inférence sur des aspects de  $s_0$ :

- $Cov(s_0(Z), S(Z)) / Var(S(Z))$   
*S est pertinent? présence d'hétérogénéité? - BLP*
- $\mathbb{E}[s_0(Z)|G]$  où  $G$  est dérivé en classant par rapport à  $S$  (groupes d'impact). *G est pertinent? Quels sont les effets sur ces groupes?*  
**GATES**
- $\mathbb{E}[Z|G]$ , caractériser les groupes en faisant un portrait de chacun  
**CLANs**



# Hétérogénéité des effets de traitement

[Chernozhukov et al., 2018]

- CATE - Conditional average treatment effect  
 $s_0(Z) = \mathbb{E}[Y(1) - Y(0)|Z]$
- Absence d'hétérogénéité :  $s_0(Z) = s_0 = \mathbb{E}[Y(1) - Y(0)]$
- Soit  $S(Z)$  un proxy de  $s_0(Z)$ , appris des données ("Generic")

[Chernozhukov et al., 2018] propose une méthode d'inférence sur des aspects de  $s_0$ :

- $Cov(s_0(Z), S(Z)) / Var(S(Z))$   
*S est pertinent? présence d'hétérogénéité? - BLP*
- $\mathbb{E}[s_0(Z)|G]$  où  $G$  est dérivé en classant par rapport à  $S$  (groupes d'impact). *G est pertinent? Quels sont les effets sur ces groupes?*  
**GATES**
- $\mathbb{E}[Z|G]$ , caractériser les groupes en faisant un portrait de chacun  
**CLANs**

# Hétérogénéité des effets de traitement

[Chernozhukov et al., 2018]

- CATE - Conditional average treatment effect  
 $s_0(Z) = \mathbb{E}[Y(1) - Y(0)|Z]$
- Absence d'hétérogénéité :  $s_0(Z) = s_0 = \mathbb{E}[Y(1) - Y(0)]$
- Soit  $S(Z)$  un proxy de  $s_0(Z)$ , appris des données ("Generic")

[Chernozhukov et al., 2018] propose une méthode d'inférence sur des aspects de  $s_0$ :

- $Cov(s_0(Z), S(Z)) / Var(S(Z))$   
*S est pertinent? présence d'hétérogénéité? - BLP*
- $\mathbb{E}[s_0(Z)|G]$  où  $G$  est dérivé en classant par rapport à  $S$  (groupes d'impact). *G est pertinent? Quels sont les effets sur ces groupes?*  
**GATES**
- $\mathbb{E}[Z|G]$ , caractériser les groupes en faisant un portrait de chacun  
**CLANs**

# Résumé de [Chernozhukov et al., 2018]

## Application pratique

Divisez le même échantillon en un échantillon principal  $M$  et un échantillon auxiliaire  $A$  à plusieurs reprises.

- Sur  $A$  : apprenez les CATEs et obtenez  $S(Z)$ , une vulnérabilité proxy au traitement.
- Sur  $M$  : confirmez les caractéristiques des CATEs (par exemple, classez les nourrissons selon  $S(Z_i)$  pour former des groupes, confirmez ou non la pertinence de ces groupes).

Agrégation quantile des inférences pour combiner les résultats entre les divisions et obtenir des tests d'hypothèses valides et des intervalles de confiance.

## Adaptation pour cet article (1/2)

- L'article original parle d'essais randomisés contrôlés → à un choc quasi-expérimental binaire  $\approx$  [Deryugina et al., 2019].
- Un grand nombre d'effets fixes non traités (municipalité) : dans la version actuelle, ces contrôles sont remplacés par un modèle détaillé de score de propension<sup>3</sup>.
- Utilisation des algorithmes de Lasso et de forêt aléatoire.
- Utilisation du package R [Welz et al., 2022].
- Groupes d'impact : les 50% les moins touchés selon le proxy, le 50-75e percentile, le 75-90e percentile, les 10% les plus touchés.
- 100 divisions pour chaque résultat d'intérêt.

---

<sup>3</sup> genre, revenu parental (introduit en déciles et linéairement), caractéristiques de la mère (âge et indicateur de naissance à l'étranger), âge gestationnel (comme indicateur de naissance prématurée et linéaire), poids à la naissance (indicateur de faible poids à la naissance et linéaire), et six autres indicateurs de santé dérivés du séjour hospitalier à la naissance

## Adaptation pour cet article (2/2)

Une nouvelle version en novembre 2022

- Les valeurs de  $p$  sont multipliées par deux avant de prendre la médiane, alors que les conseils aux praticiens ont changé depuis lors dans [Chernozhukov et al., 2022].
- Les nouvelles directives de [Chernozhukov et al., 2022] ne sont pas encore mises en oeuvre (échantillonnage stratifié, contrôle du bruit, réaction causale, etc.).

## Caractéristiques Z

Socio-economic charac.	<ul style="list-style-type: none"> <li>Gender</li> <li>Mother age</li> <li>Mother born abroad (indicator)</li> <li>Decile (indicators) and percentile (linearly) of parental income</li> <li>Benefit from the CMU-C</li> </ul>
Baseline health (from birth hospital stay)	<ul style="list-style-type: none"> <li>Birth weight (indicator of low birth and linearly)</li> <li>Gestational age (indicator of being premature and linearly)</li> <li>Healthy birth</li> <li>Normal Birth</li> <li>Stay in the neonatology department</li> <li>Respiratory or cardiovascular diagnosis</li> <li>Electrocardiogram at birth</li> </ul>
Local charac.	<ul style="list-style-type: none"> <li>Radiography of the respiratory system at birth</li> <li>PM2.5 exposure</li> <li>Accessibility to GP</li> <li>Accessibility to paediatrician</li> <li>Type/size of the urban area (6 modalities)</li> </ul>

Table: List of Z features

## Présence d'hétérogénéité?

$$\beta_2 = \text{Cov}(s_0(Z), S(Z)) / \text{Var}(S(Z)) = 0?$$

	Intermediate shock of exposure $\bar{n} = 7$	Large shock of exposure $\bar{n} = 10$
<i>Drug delivery for obstructive airway diseases</i>		
- before first anniversary		
$\hat{\beta}_2$	0.105 [0.028;0.181] p=0.007	0.114 [0.020;0.208] p=0.018
- between first and third anniversary		
$\hat{\beta}_2$	0.207 [-0.141;0.575] p=0.243	0.330 [-0.064;0.718] p=0.097
<i>Hospital emergency admissions</i>		
- asthma before third anniversary		
$\hat{\beta}_2$	0.167 [0.069;0.265] p=0.001	0.156 [0.021;0.287] p=0.020
- bronchiolitis before second anniversary		
$\hat{\beta}_2$	0.330 [0.253;0.408] p=0.000	0.282 [0.188;0.377] p=0.000
- bronchiolitis or asthma before third anniversary		
$\hat{\beta}_2$	0.184 [0.084;0.278] p=0.000	0.122 [-0.013;0.258] p=0.073

## Un groupe d'impact ressort-il?

---



---

### *Drug delivery for obstructive airway diseases before first anniversary*

$\hat{\gamma}_1$	-0.002 [-0.012;0.009]	$p=0.761$
$\hat{\gamma}_2$	0.007 [-0.007;0.022]	$p=0.324$
$\hat{\gamma}_3$	0.009 [-0.010;0.028]	$p=0.324$
$\hat{\gamma}_4$	0.023 [-0.001;0.046]	$p=0.054$
$\hat{\gamma}_4 - \hat{\gamma}_1$	0.024 [-0.001;0.049]	$p=0.059$

---

### *Hospital emergency admissions*

#### Bronchiolitis before second anniversary

$\hat{\gamma}_1$	-0.002 [-0.006;0.003]	$p=0.495$
$\hat{\gamma}_2$	0.000 [-0.006;0.006]	$p=0.973$
$\hat{\gamma}_3$	-0.000 [-0.009;0.008]	$p=0.988$
$\hat{\gamma}_4$	0.015 [0.004;0.026]	$p=0.005$
$\hat{\gamma}_4 - \hat{\gamma}_1$	0.016 [0.005;0.028]	$p=0.005$

---



---

Evidence of two groups of 10% mostly affected.



# Caractéristiques des 10% les plus affectés

	Average of the characteristics for :					
	(i) the 10% the most affected,			(ii) the 50% the least affected, and (iii) difference most - least		
	<i>Bronchiolitis emergency before second anniversary</i>			<i>Antiasthmatic medication before first anniversary</i>		
	(i)	(ii)	(i)-(ii)	(i)	(ii)	(i)-(ii)
Girl	0.463	0.476	-0.012 (p=0.020)	0.426	0.490	-0.062 (p=0.000)
Birth weight	3082	3284	-203 (p=0.000)	3231	3275	-45 (p=0.000)
Gestational age	38.22	39.09	-0.866 (p=0.000)	38.88	39.02	-0.148 (p=0.000)
No pathology	0.693	0.831	-0.140 (p=0.000)	0.803	0.822	-0.019 (p=0.000)
Cardio-respiratory pathology	0.251	0.165	0.085 (p=0.000)	0.177	0.170	0.007 (p=0.080)
Monitoring electrocardiogram	0.121	0.038	0.081 (p=0.000)	0.058	0.044	0.014 (p=0.000)
Premature birth	0.281	0.115	0.165 (p=0.000)	0.166	0.126	0.041 (p=0.000)
Low birth weight	0.196	0.056	0.140 (p=0.000)	0.085	0.070	0.014 (p=0.000)
Birth without significant issues	0.576	0.741	-0.164 (p=0.000)	0.692	0.732	-0.040 (p=0.000)
Disposable income percentile	43.90	49.62	-5.586 (p=0.000)	48.76	50.18	-1.406 (p=0.000)
Neonatology stay	0.230	0.082	0.148 (p=0.000)	0.113	0.095	0.019 (p=0.000)
Respiratory system radiography	0.086	0.029	0.057 (p=0.000)	0.041	0.034	0.008 (p=0.000)
Accessibility : GP	3.99	4.10	-0.113 (p=0.000)	4.03	4.10	-0.072 (p=0.000)
Accessibility : Paediatrician	3.79	3.45	0.334 (p=0.000)	3.40	3.56	-0.148 (p=0.000)
Decile 1	0.176	0.089	0.087 (p=0.000)	0.114	0.093	0.019 (p=0.000)
Decile 10	0.106	0.085	0.020 (p=0.000)	0.093	0.101	-0.007 (p=0.018)
CMU beneficiary	0.173	0.133	0.040 (p=0.000)	0.126	0.125	-0.000 (p=0.829)
PM2.5	13.12	11.18	1.94 (p=0.000)	11.74	11.51	0.221 (p=0.000)
Mother Age	30.45	30.62	-0.182 (p=0.002)	30.10	30.84	-0.761 (p=0.000)
Mother Born abroad	0.245	0.163	0.084 (p=0.000)	0.173	0.164	0.009 (p=0.017)
Outside Cities	0.062	0.052	0.009 (p=0.000)	0.059	0.054	0.003 (p=0.087)
Urban Area <50,000	0.091	0.112	-0.022 (p=0.000)	0.112	0.108	0.003 (p=0.199)
Urban Area 50,000 - 200,000	0.130	0.186	-0.056 (p=0.000)	0.178	0.171	0.008 (p=0.045)
Urban Area 200,000 - 700,000	0.177	0.230	-0.053 (p=0.000)	0.206	0.223	-0.017 (p=0.000)
Urban Area >700,000	0.184	0.200	-0.016 (p=0.000)	0.201	0.210	-0.009 (p=0.028)
Urban Area of Paris	0.352	0.218	0.130 (p=0.000)	0.239	0.233	0.005 (p=0.116)

## Récapitulatif

- Sous l'hypothèse de restriction d'exclusion, il existe un lien causal entre la santé respiratoire des enfants de moins de 3 ans et la pollution de l'air.
- Selon la méthode de [Chernozhukov et al., 2018], les effets sont très hétérogènes (10% des enfants sont affectés, avec des impacts importants).
- Ceux qui sont le plus touchés ont de mauvais indicateurs de santé de base et sont plus exposés. Ils font également le plus souvent partie des 10% des enfants les plus pauvres (surtout pour les admissions aux urgences).
- Ceux qui sont le plus touchés en termes de consommation de médicaments et en termes d'admissions aux urgences sont différents : quel est le rôle de l'accès aux soins de santé ?

## Discussion

- Absence d'instrumentation (disponibilité des données, disparités du PM2,5 entre les sources)
- Un choix de seuil pour équilibrer la taille du groupe traité et la taille du choc de pollution de l'air
- Une spécification binaire pour mener l'analyse de l'hétérogénéité
- Effets probablement sous-estimés lorsqu'on se concentre sur la consommation de soins de santé plutôt que sur la santé observée

Questions? Merci!

# Bibliography I



Arceo, E., Hanna, R., and Oliva, P. (2016).

Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city.  
*The Economic Journal*, 126(591):257–280.



Bind, M.-A. (2019).

Causal modeling in environmental health.  
*Annual review of public health*, 40:23–43.



Breda, T., Grenet, J., Monnet, M., and Van Effenterre, C. (2018).

*Can Female Role Models Reduce the Gender Gap in Science?: Evidence from Classroom Interventions in French High Schools.*  
Paris-Jourdan Sciences Economiques.



Chernozhukov, V., Demirer, M., Duflo, E., et al. (2022).

Generic machine learning inference on heterogenous treatment effects in randomized experiments.  
Technical report, arXiv.org.



Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2018).

Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india.  
Technical report, National Bureau of Economic Research.



Currie, J. (2011).

Inequality at birth: Some causes and consequences.  
*American Economic Review*, 101(3):1–22.

## Bibliography II



Dechezleprêtre, A., Rivers, N., and Stadler, B. (2019).  
The economic cost of air pollution: Evidence from europe.



Deguen, S. and Zmirou-Navier, D. (2010).  
Social inequalities resulting from health risks related to ambient air quality—a european review.  
*European Journal of Public Health*, 20(1):27–35.



Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019).  
The mortality and medical costs of air pollution: Evidence from changes in wind direction.  
*American Economic Review*, 109(12):4178–4219.



Dominici, F., Greenstone, M., and Sunstein, C. R. (2014).  
Particulate matter matters.  
*Science*, 344(6181):257–259.



Godzinski, A. and Suarez Castillo, M. (2021).  
Disentangling the effects of air pollutants with many instruments.  
*Journal of Environmental Economics and Management*, 109:102489.



Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C.,  
Garay, M. J., Kalashnikova, O. V., Kahn, R. A., et al. (2020).  
Global estimates and long-term trends of fine particulate matter concentrations (1998–2018).  
*Environmental Science & Technology*, 54(13):7879–7890.

# Bibliography III



Imai, K. and Li, M. L. (2022).

Statistical inference for heterogeneous treatment effects discovered by generic machine learning in randomized experiments.

*arXiv preprint arXiv:2203.14511.*



Isen, A., Rossin-Slater, M., and Walker, W. R. (2017).

Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970.

*Journal of Political Economy*, 125(3):848–902.



Jans, J., Johansson, P., and Nilsson, J. P. (2018).

Economic status, air quality, and child health: Evidence from inversion episodes.

*Journal of health economics*, 61:220–232.



Real, E., Couvidat, F., Ung, A., Malherbe, L., Raux, B., Gressent, A., and Colette, A. (2022).

Historical reconstruction of background air pollution over France for 2000–2015.

*Earth System Science Data*, 14(5):2419–2443.



Schwartz, J., Bind, M.-A., and Koutrakis, P. (2016).

Estimating causal effects of local air pollution on daily deaths: effect of low levels.

*Environmental health perspectives*, 125(1):23–29.



Welz, M., Alfons, A., Demirev, M., and Chernozhukov, V. (2022).

Genericml: Generic machine learning inference. R package version 0.2.1.

## PM2.5

Source: PM2.5 de Surface, V4.EU.03 2001-2018, produit par l'Atmospheric Composition Analysis Group<sup>4</sup> [Hammer et al., 2020]

- Mesures satellites + algorithmes de reconstitution de l'épaisseur optique des aérosols
- Modèle de chimie transport GEOS-chem (Réanalyse MERRA-II): relation PM2.5 en surface, AOD
- Traitement statistique final (GWR) pour calibrer sur stations de mesure

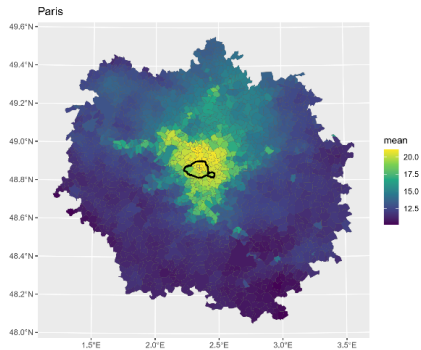
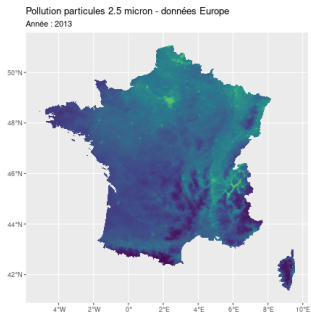
Données publiques, diffusées à haute résolution ( $0.01 \times 0.01^\circ$ ) pour réagrégation ad-hoc, en moyenne annuelle; ici réagrégée à la commune

---

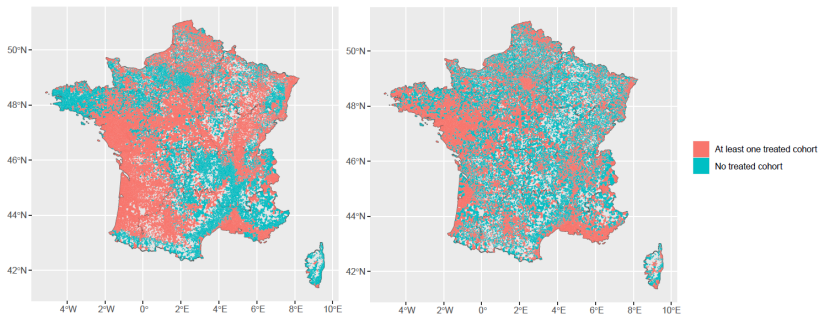
<sup>4</sup>Washington university in St Louis



# PM2.5



## Treatment spatial repartition



**Figure:** Left: Municipality with an absolute high number of thermal inversion, at least a year over 2008-2017. Right: Municipality with at least one treated cohort, or no treated cohort ( $\bar{n} = 7$ )

Back

## Characteristics of treated and non treated infants [Back](#)

	N	Mean values		Conditional Equality Test	
		Treated	Non treated	Estimate	p-value
<i>Treatment definition</i>					
Days with thermal inversion in first year	336, 169	50.9	32.8	10.8	0.00
<i>Birth hospital stay</i>					
Missing birth hospital stay	336, 169	0.13	0.15	-0.00	0.42
Birth weight	287, 517	3252.6	3260.0	-1.01	0.80
Gestational age	266, 795	39.0	39.0	0.03	0.09
Healthy birth (absence of diagnostic)	287, 518	0.83	0.81	0.00	0.58
Duration	287, 518	4.9	5.0	-0.01	0.79
With respiratory or cardiovascular affections	287, 518	0.18	0.17	0.00	0.44
Electrocardiogram	287, 518	0.05	0.05	0.00	0.69
Respiratory system radiography	287, 518	0.04	0.04	-0.00	0.36
<i>Household characteristics (tax data)</i>					
Missing household disposable income	336, 169	0.13	0.19	0.00	0.62
Percentile of household disposable income	275, 054	49.3	49.6	0.05	0.79
Household number of persons	199, 942	4.0	4.0	-0.01	0.54
Household number of dependants	199, 981	1.9	2.0	-0.01	0.42
Household below poverty line	200, 103	0.10	0.10	0.01	0.03
Household owns its housing	200, 103	0.49	0.48	-0.00	0.50
<i>Household characteristics (Civil registry)</i>					
Girl	336, 169	0.49	0.49	0.00	0.44
Mother age	335, 719	30.4	30.5	-0.11	0.00
Mother born abroad	334, 865	0.18	0.20	-0.00	0.76
Father born abroad	330, 367	0.19	0.21	0.00	0.54
Mother born abroad, no French nationality	332, 666	0.12	0.14	-0.00	0.35
Father born abroad, no French nationality	328, 403	0.13	0.14	-0.00	0.94

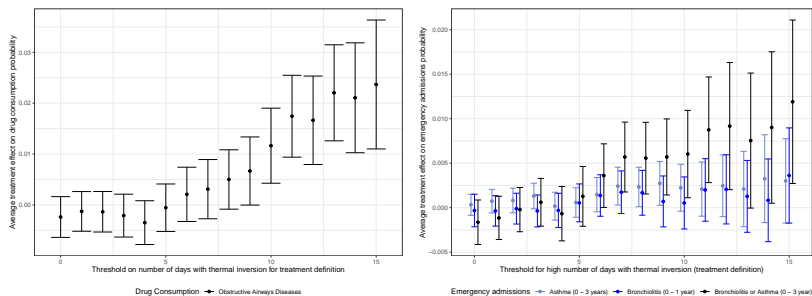
# Binary shock of exposure to air pollution [Back](#)

	Intermediate shock of exposure $\bar{n} = 7$		Large shock of exposure $\bar{n} = 10$	
<i>Additional days with thermal inversion</i>				
	10.79	10.60	11.38	11.23
	[10.48;11.11]	[10.32;10.89]	[10.92;11.83]	[10.76;11.71]
$X^I$	no	yes	no	yes
Sample size	336169	217859	336169	217859
<i>Additional exposure to PM2.5 in first year, by source, for those born in January</i>				
ACAG (2008-2017)				
	0.12	0.055	0.23	0.14
	[0.022;0.22]	[-0.044;0.15]	[0.11;0.35]	[-0.014;0.30]
$X^I$	no	yes	no	yes
Sample size	82164	49703	82164	49703
Ineris (2009-2017)				
	0.24	0.23	0.37	0.36
	[0.12;0.37]	[0.085;0.37]	[0.17;0.56]	[0.13;0.60]
$X^I$	no	yes	no	yes
Sample size	73354	49214	73354	49214

# Characteristics of the sample and birth comparability across treated and non treated infants

	N	Mean values		Conditional Equality Test	
		Treated	Non treated	Estimate p-value	
<i>Treatment definition</i>					
Days with thermal inversion in first year	336,169	50.9	32.8	10.8 0.00	
Outside Cities	336,169	0.06	0.05		
Urban Area <50,000	336,169	0.12	0.10		
Urban Area 50,000 - 200,000	336,169	0.19	0.16		
Urban Area 200,000 - 700,000	336,169	0.23	0.21		
Urban Area >700,000	336,169	0.24	0.21		
Urban Area of Paris	336,169	0.16	0.26		
Born in 2012 or before	336,169	0.35	0.54		
Born after 2012	336,169	0.65	0.46		

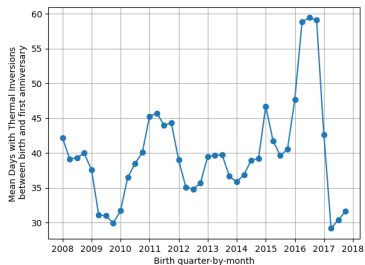
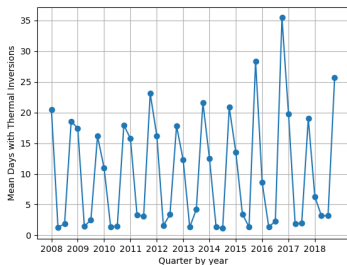
# Health outcomes and unusually high number of days with thermal inversions

[Back](#)


**Figure:** Days with Thermal inversion and Health Outcomes, threshold

$$Y_i = \beta \times 1 \left\{ N_i - \hat{N}_i \geq \bar{n} \right\} + X_i^W \gamma_W + \mu_c + \delta_t + \epsilon_i,$$

# Thermal Inversion Temporal Variations



**Figure:** Temporal variations in the number of days with thermal inversions quarter by quarter, and in total the next 12 months