Les Inégalités sociales de santé liées aux effets de la pollution sur la santé

Emmanuelle Lavaine

UNIVERSITÉ PARIS 1 - PANTHÉON SORBONNE

U.F.R DE SCIENCES ECONOMIQUES

Année 2012-2013 Numéro attribué par la bibliothèque

| | | | | | | | | | | | | |

Thèse pour le Doctorat de Sciences Economiques

soutenue publiquement par

Emmanuelle LAVAINE

12 Septembre 2013

LES INÉGALITÉS SOCIALES LIÉES AUX EFFETS DE LA POLLUTION SUR LA SANTÉ

Directeur de thèse

Madame Mireille Chiroleu-Assouline - Professeur à l'Université Paris 1 Panthéon-Sorbonne

Madame Carmen Arguedas Tomás - Professeur à l'Université Autonoma de Madrid

Jury

Monsieur Olivier Chanel - Directeur de recherche - CNRS GREQAM

Monsieur James Hammit - Professeur à Toulouse School of Economics et Harvard

Madame Katheline Schubert - Professeur à l'Université Paris 1 Panthéon-Sorbonne

Madame María Teresa Blázquez Cuesta - Professeur à l'Université Autonoma de Madrid

Avertissement

Mis à part l'introduction et la conclusion de cette thèse, les différents chapitres sont issus d'articles de recherche rédigés en anglais et dont la structure est autonome. Par conséquent, des termes "papier" ou "article" y font référence, et certaines informations, notamment la littérature, sont répétées d'un chapitre à l'autre.

Notice

Except the general introduction and the conclusion, all chapters of this thesis are self-containing research articles. Consequently, terms "paper" or "article" are frequently used. Moreover, some explanations, like corresponding literature, are repeated in different places of the thesis.

Remerciements

J'adresse de très chaleureux remerciements à ma directrice de thèse Mireille Chiroleu-Assouline. Elle m'a beaucoup soutenue et orientée sur des sujets pertinents. Nos rendez-vous hebdomadaires du mercredi durant ces quelques années m'ont apporté une aide et un enseignement précieux. Elle m'a indéniablement transmis son goût pour la recherche. Je lui dois beaucoup dans la réussite de cette thèse. Je la remercie pour sa grande disponibilité, sa confiance et sa gentillesse. Elle est devenue bien plus qu'une directrice de thèse.
Je tiens également à remercier ma co-directrice Carmen Arguedas Tomás pour ses bons conseils, son soutien. Sa co-direction m'a permis d'intégrer quelques temps un autre laboratoire de recherche et de faire partie d'un programme de recherche espagnol. J'adresse également mes remerciements à Olivier Chanel, James Hammit, Katheline Schubert et María Teresa Blázquez Cuesta pour avoir accepté d'être les rapporteurs de cette thèse. J'ai incorporé le plus minutieusement possible leurs commentaires, qui je suis sûre ont fortement amélioré la qualité de la thèse.

Je tiens également à remercier très chaleureusement Jean De Beir pour m'avoir mise sur le chemin du doctorat et de la Maison des Sciences Economiques (MSE). J'ai débuté mes années de thèse à la MSE où l'atmosphère et l'environnement sont très propices à la recherche. On y croise toujours quelqu'un pour vous prodiguer un bon conseil, vous rappeler l'heure d'un séminaire ou simplement pour discuter autour d'un café. Je veux tout d'abord remercier Jean-Olivier Hairault, le directeur de l'axe. Je remercie aussi Katrin Millock, Antoine d'Autume, Hippolyte d'Albis, Nicolas Jacquemet, Mouez Fodha pour leurs conseils. Je pense aussi à l'ensemble des doctorants de la MSE et particulièrement aux doctorants et postdoc du troisième: Camille et Guillaume avec qui j'ai débuté ma thèse et découvert les rouages du statut de doctorant; Esther, Djamel, Pablo, Cheng, Claire, Basile, Axelle et Diana que j'ai rencontrés par la suite et avec qui j'ai pris baucoup de plaisir à discuter, échanger, partager les pauses déjeuner, et aux nouveaux doctorants de mon bureau et de celui d'à côté : Barish, Estefania, Anastasia, Hamze, Lorenzo, Vincent, Antoine pour leur sympathie. Merci aux nombreux doctorants du 4ème et un merci particulier à Noémi pour avoir partagé nos petits soucis de thésards. Les conférences et séjours de recherche ont aussi été l'occasion de nombreuses rencontres. Je remercie Matthew Neidell pour m'avoir fait confiance et m'avoir inclue dans son département de recherche aux Etats-Unis. J'ai beaucoup appris à ses côtés lors de notre collaboration sur le deuxième chapitre de cette thèse. Je veux remercier Sanja Pekovic pour m'avoir fait découvrir la recherche via Skype, Carlos Llano pour sa bienvenue à Madrid, Lisa Anouliès pour le partage de ses expériences. Mes remerciements s'adressent également au laboratoire d'Economie de Toulouse pour m'avoir accueillie dans ma dernière année de thèse et particulièrement François Salanié, Nicolas Treich, Sylvain Chabbé Ferret, Enrik Andersson, Alban Thomas, Stephan Ambec pour leur aide et leurs conseils. J'exprime également mes remerciements à "la troupe" des économistes de l'environnement que je prends plaisir à retrouver à chaque conférence Alain Ayong le cama, Vincent Martinet, Saraly Andrade, Julien Daubannes et bien d'autres. Je remercie les doctorants de Toulouse, en particulier Yann, Tunch et Ananya, et les doctorants de la Autonoma de Madrid, José, Tamara et Nuria, sans oublier les doctorants de Columbia qui m'ont accueillie très chaleureusement. Mes remerciements s'adressent aussi à l'université Paris 1 pour le financement de mon travail de thèse et ma mobilité de co-tutelle de thèse en Espagne ainsi que les responsables du programme Alliance Paris 1 Sciences po Polytechnique pour le financement de mon séjour doctoral à l'université de Columbia. Merci à Alessia Lefebure et Lauranne Bardin d'en avoir facilité mon séjour.

Cette thèse utilise diverses bases de données riches en informations. J'exprime donc mes remerciements à Souad Boualala de l'Ademe, Laurence Rouil et Maxime Beauchamp de l'Ineris pour la diffusion et le suivi des données de pollution. Je remercie Morgan Hamon et Florence Mirrione pour la diffusion et l'aide à la compréhension des bases du PMSI, Emilie Masson et Florence Celen pour les procédures CNIL, Isabel Giagnoni et Adélia Nobre pour l'aide au ciblage de mes besoins et la diffusion de la base PERVAL. Je remercie aussi Veronique Malet-Ferroni pour son aide administrative dans l'obtention des données et le montage des dossiers. Evidemment, je remercie vivement Viviane Makougni, Elda André, Loïc Sorel et Eric Zyla pour leur efficacité, leur gentillesse et leur aide continuelle pendant ces années de thèse.

Merci à tous ceux que j'ai rencontrés pendant ces années de thèse et dont j'aurai omis d'adresser mes remerciements. Enfin, je voudrais remercier ma soeur, mon frère et mes parents pour leur soutien infaillible. Je leur dédie cette thèse. Merci à Pierre d'avoir écouté longuement mes innombrables présentations et de m'avoir épaulée.

Ces années de doctorats furent rythmées par de multiples rencontres avec les chercheurs, professeurs, doctorants et des professionnels de tout horizon qui ont enrichi mon travail. La découverte de métiers passionnants liés à la recherche a confirmé mon désir d'évoluer dans cet environnement.

Contents

1. This chapter is a joint work with Matthew Neidell from Columbia University

List of Figures

List of Tables

Chapter 1

Introduction générale

Les émissions de polluants atmosphériques ont tendance à diminuer dans l'ensemble des pays industrialisés. Néanmoins, l'exposition actuelle à la pollution atmosphérique engendre encore des effets néfastes sur la santé. L'OMS estime que la pollution de l'air est à l'origine d'environ deux millions de décès prématurés par an à l'échelle mondiale. Dans ce contexte, l'impact de la pollution environnementale sur la santé est une priorité croissante de santé publique. Mon projet de thèse consiste donc à étudier les impacts sanitaires liés à la dégradation de l'environnement et d'analyser leurs conséquences macroéconomiques.

1.1. De l'impact de la pollution sur la santé aux disparités environnementales...

1.1.1. L'impact de la pollution atmosphérique sur la santé en France

La pollution atmosphérique est un mélange de nombreuses substances connues et inconnues. En France, certaines sont soumises à la réglementation et/ou font l'objet d'une surveillance de la part des réseaux de mesure de la qualité de l'air: les oxydes d'azote (NOx), le dioxyde de soufre (SO_2), les particules fines (PM), les hydrocarbures appelés Composés Organiques Volatils Non Méthaniques (COVNM) ainsi que les trois principaux gaz à effet de serre : le dioxyde de carbone (CO_2), le méthane (CH_4) ainsi que le protoxyde d'azote (N2O). Ces polluants entraînent une réaction inflammatoire des bronches, favorisent les pathologies préexistantes et jouent un rôle de cofacteur dans la réaction allergique car ils augmentent la réponse des bronches aux effets des allergènes. Une étude française par des chercheurs de l'Inserm montre qu'une personne asthmatique déclenchera plus facilement une crise ou sera plus gênée lors d'un pic de pollution (Yamada *et al.*, 2002). Pour reprendre Chanel *et al.* (2000), il est difficile d'associer un effet sanitaire spécifique à un indicateur de pollution particulier. Néanmoins, la littérature épidémiologique travaille à définir l'impact de chacun de ces polluants sur la santé humaine.

Tout d'abord, les particules en suspension classées en fonction de leur diamètre aérodynamique: PM_{10} (diamètre de moins de 10 μm); $PM_{2.5}$ (diamètre de moins de 2,5 μm) présentent des effets sur la santé plus importants que tout autre polluant. Les particules $PM_{2.5}$ sont les plus dangereuses car, après inhalation, elles peuvent atteindre la région alvéolaire et altérer les échanges gazeux à l'intérieur des poumons. Cette pollution entraîne des maladies respiratoires obstructives chroniques et le cancer du poumon chez l'adulte (World Health Organization (WHO), 2006). Des projets européens associent les particules en suspension à la mortalité ou la morbidité tel NMMAPS (Peng *et al.*, 2004) ou APHEA (Katsouyanni *et al.*, 1995), ainsi que de nombreuses études épidémiologiques récentes (Schikowski *et al.*, 2008) , (Janke *et al.*, 2009) , (Annesi-Maesano *et al.*, 2007)

. Les particules en suspension seraient responsables de 40 000 morts par an selon une estimation pour l'Autriche, la France et la Suisse (Kunzli *et al.*, 2000). On sait aussi que la pollution de l'air à l'intérieur des habitations due à l'utilisation de combustibles solides est responsable de 1,6 million de décès dans le monde selon une note récente de l'OMS (World Health Organization (WHO), 2009). Or, les particules comme les oxydes de soufre font partie des polluants dégagés par l'utilisation de ces combustibles. Les particules en suspension souvent émises en même temps que le SO_2 en potentialisent les effets (Lecoq *et al.*, 2009).

Le dioxyde de soufre (SO_2), un des polluants majeurs émis par les raffineries de pétrole est le principal polluant du secteur industriel. Le SO_2 est responsable d'un certain nombre de problèmes respiratoires, il est donc réglementé aux Etats-Unis comme en Europe. En Europe, le seuil pour le dioxyde de soufre est fixé par la norme européenne 2002-2013 à $350\mu g/m^3$ en moyenne horaire pour protéger la santé humaine. En comparaison, depuis juin 2010, la réglementation aux Etats-Unis fixe le seuil en moyenne horaire à 75 parties par milliards soit environ 175 $\mu g/m^3$.[1]

Le SO_2 est un gaz incolore, soluble dans l'eau, qui est réactif et très irritant pour les yeux et les voies respiratoires. Le SO_2 est soumis à une série de processus de transformation dans l'atmosphère. Il se transforme en acide sulfurique mais peut aussi réagir avec d'autres composés dans l'atmosphère pour former de petites particules. Ces particules pénètrent profondément dans les parties sensibles des poumons et peuvent causer ou aggraver des maladies respiratoires, telles que l'emphysème et la bronchite, et aggraver une maladie

1. ppb est un volume de gaz polluant par 10^9 de volume d'air. 1 ppb = 2.66 μg/m^3 pour le SO_2

cardiaque préexistante, conduisant à une augmentation des hospitalisations et des décès prématurés.

Le SO_2 représente donc un risque environnemental majeur pour la santé et son rôle délétère a été clairement démontré dans de nombreuses études. Ce gaz est associé à une altération de la fonction pulmonaire chez l'enfant, à une exacerbation des symptômes respiratoires aigus chez l'adulte, à des crises d'asthme (Laurent *et al.*, 2007) et provoque des irritations oculaires (Lecoq *et al.*, 2009). Le nombre d'admissions à l'hôpital pour des cardiopathies et la mortalité augmentent à la suite de journées avec de fortes concentrations en SO_2 (Finkelstein *et al.*, 2003). Selon l'Institut de Veille Sanitaire (IVS), une exposition au SO_2 sur de longues périodes à des niveaux moyens journaliers faibles est significativement corrélée à la survenue de différents événements sanitaires tels que l'hospitalisation pour maladies respiratoires et cardio-vasculaires ischémiques, l'aggravation des insuffisances respiratoires chroniques et les décès pour pathologies cardio-vasculaires (Lecoq *et al.*, 2009).

Les gaz à effet de serre, principalement le gaz carbonique (CO_2) qui est un indicateur de la consommation énergétique car une conséquence de la combustion d'énergies fossiles, sont reconnus dangereux pour la santé publique par l'agence américaine de protection de l'environnement (Environmental Protection Agency (EPA), 2011). Cette dernière a conclu que les gaz à effet de serre contribuaient au changement climatique et donc à la pollution de l'air. Cette étude démontre que le changement climatique conduit à des concentrations plus élevées en ozone au niveau du sol.

Or, l'ozone, directement associé à la pollution, est un polluant à propriétés oxydantes, dangereux pour la santé. Ce gaz nocif est à l'origine d'irritations des muqueuses oculaires et respiratoires et de crises d'asthme chez les sujets sensibles. À des concentrations trop élevées, l'ozone a des effets marqués sur la santé de l'homme. On observe alors des problèmes respiratoires, le déclenchement de crises d'asthme, une diminution de la fonction pulmonaire et l'apparition de maladies respiratoires (World Health Organization (WHO), 2009). A l'identique, l'Agence de protection environnementale des Etats-Unis stipule que l'ozone aggrave l'asthme, provoque la toux et l'irritation, une inflammation du tissu respiratoire et une plus grande susceptibilité d'infection respiratoire (Environmental Protection Agency (EPA), 2011). Des études épidémiologiques récentes abondent dans le même sens et mettent en exergue que l'ozone est associé à de forts taux de mortalité (Janke *et al.*, 2009) et à l'exacerbation de l'asthme (Neidell, 2009), (Laurent *et al.*, 2007), (Wilhelm *et al.*, 2009). Il existe cependant peu d'évidence d'un effet de long terme de l'ozone sur le cancer ou la mortalité (World Health Organization (WHO), 2006).

Le dioxyde d'azote (NO_2) est un oxydant puissant qui pénètre dans les voies aériennes inférieures notamment les bronchioles. Pour des expositions élevées, on constate des lésions inflammatoires de l'épithélium, une augmentation des lymphocytes et des macrophages au niveau cellulaire (Lecoq *et al.*, 2009). Un certain nombre d'études épidémiologiques ont aussi permis d'évaluer les effets de court terme de la pollution par le NO_2 sur la fonction respiratoire (Schikowski *et al.*, 2008), (Laurent *et al.*, 2007). L'augmentation des niveaux de NO_2 est aussi corrélée à une augmentation de la mortalité (Janke *et al.*, 2009) et des hospitalisations pour pathologies respiratoires (Annesi-Maesano *et al.*, 2007). Certaines études expérimentales ont montré un rôle facilitateur joué par le dioxyde d'azote et des particules de diesel dans la sensibilisation allergénique.

Les hydrocarbures organiques volatils tels le benzène, le toluène, le xylène et l'éthylbenzène ont aussi des effets importants sur la santé. Les effets de l'inhalation de vapeurs de xylène ou de toluène sont la dépression du système nerveux central (SNC), des lésions hépatiques et rénales réversibles et des symptômes d'oedème pulmonaire. Les effets de l'éthylbenzène sur la santé humaine ne sont pas encore totalement compris. Cependant, en se fondant sur des preuves encore suffisantes chez les animaux de laboratoire et des preuves insuffisantes chez les humains, le Centre International de Recherche sur le Cancer a conclu que l'éthylbenzène est peut-être cancérogène chez les humains (Midorikawa *et al.*, 2004).

La pollution de l'air représente un risque environnemental majeur pour la santé et son rôle délétère a été clairement démontré dans de nombreuses études entraînant crise et exacerbation de l'asthme, maladies respiratoires obstructives, maladies cardio-vasculaires et cancer du poumon. Cependant, les données disponibles sur les effets de la pollution de l'air sont insuffisantes, particulièrement en ce qui concerne les effets de long terme.

1.1.2. Les inégalités environnementales

Les recherches portant sur les inégalités sociales en matière d'environnement sont encore récentes en Europe et très limitées en France (Viel *et al.*, 2011). Des inégalités sociales en matière d'exposition aux risques environnementaux ont été observées dans de nombreux pays industrialisés justifiant l'importance de continuer à étudier ce sujet (Laurent *et al.*, 2007). Cette thèse s'inscrit dans un nouvel axe de recherche sur le concept de justice environnementale et fournit une vue d'ensemble sur la distribution des risques environnementaux. La littérature rend compte de la corrélation négative entre l'exposition à la pollution et le statut socioéconomique à travers le taux de chômage et l'éducation au Canada (Premji *et al.*, 2007), les groupes ethniques [2] et le revenu en Angleterre (McLeod *et al.*, 2000) et aux Etats-Unis (Grineski *et al.*, 2007), (Morello-Frosch *et al.*, 2002) où le concept de justice environnementale est l'objet d'une attention croissante.

Le concept de justice environnementale largement utilisé aux Etats-Unis est une notion encore nouvelle en Europe (Laurent, 2011). La notion de justice environnementale intervient en 1994 aux Etats-Unis par le biais de l'agence américaine de protection de l'environnement (EPA) comme "le traitement équitable des gens de toutes races, cultures et revenus dans le développement des règlements, lois et politiques environnementales". Depuis cette date, les projets d'implantation tels que des installations de sites industriels polluants ou de traitement de déchets, doivent respecter une certaine équité environnementale. Peu d'études européennes examinent l'importance des facteurs socio-économiques sur les liens entre pollution de l'air et santé (Deguen & Zmirou, 2012).

Pourtant en France, des territoires comme le Nord-Pas de Calais ou les Zones urbaines sensibles (ZUS) peuvent présenter des cumuls d'inégalités avec des populations défavorisées qui vivent dans des environnements altérés. Nous connaissons néanmoins peu de choses sur les différences d'exposition entre population. Ces disparités environnementales peuvent être liées aux risques naturels et techniques, à la présence d'une raffinerie sur le territoire,

2. Les données sur l'ethnie ne sont pas disponibles en France et ne permettent donc pas d'étudier les inégalités en France de la même manière qu'elles sont étudiées aux Etats-Unis.

ou à la qualité du cadre de vie (bruit), et à des aménités environnementales (espaces verts, forêts). Ces différences peuvent se retrouver à une vaste échelle géographique comme le département ou à une échelle beaucoup plus fine comme celle de la commune. Il reste cependant difficile de tracer une frontière entre la dimension sociale des inégalités liées au revenu, à la catégorie socio-professionnelle, aux minorités éthniques, à l'éducation et à la spécificité de la zone géographique.

En France, les sources de pollution environnementale telles que les sites SEVESO seuils hauts sont répartis sur l'ensemble du territoire. La directive dite Seveso ou directive 96/82/CE est une directive européenne qui impose aux états membres de l'Union Européenne d'identifier les sites industriels présentant des risques d'accidents majeurs. Même si le nombre de sites a tendance à se réduire suite à l'évolution perpétuelle de la norme SEVESO[3], de nombreux sites restent encore présents. Néanmoins, ces sites sont principalement associés aux grands bassins industriels récents ou passés: étang de Berre, les bassins miniers du nord, la région Ile-de-France, la région lyonnaise, etc. L'examen géographique des sources de pollution environnementale nous montre que celles-ci sont situées principalement dans les zones industrielles.

Selon Laigle & Oehler (2004), quatre aspects complémentaires doivent être ainsi envisagés pour caractériser les inégalités écologiques : tout d'abord, les inégalités territoriales : elles résultent à la fois de disparités naturelles mais également des facteurs économiques, politiques et sociaux qui ont façonné les territoires ; il existe également des inégalités d'accès à l'urbanité et à la qualité du cadre de vie puis des inégalités dans la capacité d'agir sur

3. A compter du 1er juin 2015, de nouvelles exigences seront applicables aux établissements afin de prévenir et de mieux gérer les accidents majeurs impliquant des produits chimiques dangereux.

l'environnement et d'interpeller la puissance publique pour transformer le cadre de vie. Finalement, il existe aussi des inégalités d'exposition aux nuisances urbaines et aux risques (naturels, technologiques, industriels) que je m'attache à développer plus particulièrement dans cette thèse.

1.2. Problématique

Ce projet de thèse consiste à étudier les inégalités sociales de santé liées à la dégradation de l'environnement et d'analyser leurs conséquences macroéconomiques. Le but est d'aboutir à la mise en évidence de différences d'exposition, mais aussi de conséquences en termes de mortalité, de morbidité, et donc de pertes de productivité, entre catégories de ménages, selon leur lieu d'habitation, leur revenu, etc. Cette thèse s'inscrit dans un nouvel axe de recherche sur le concept de justice environnementale et fournit une vue d'ensemble sur la distribution des risques environnementaux.

Malgré une tendance générale à la baisse des émissions, la France ne respecte pas toutes les normes européennes (pour les $PMSO_{10}$ et les NOx notamment) et des progrès restent donc à faire au niveau de l'amélioration de la qualité de l'air. Par ailleurs, pour répondre à une demande énergétique internationale croissante, le pétrole brut est traité et transformé, le but étant de tirer le maximum de produits à haute valeur commerciale. Les processus de raffinage posent, néanmoins, un risque sanitaire. Les liens entre santé et pollution ont été fortement relatés dans la littérature internationale en se focalisant sur les taux de mortalité en Autriche, France et Suisse (Kunzli *et al.*, 2000), en Angleterre (Janke *et al.*, 2009), sur la sensibilisation allergique en France (Annesi-Maesano *et al.*, 2007), sur l'asthme (Wilhelm et al. 2009), ou sur les risques de cancer aux Etats-Unis (Chay *et al.*, 2003), (Morello-Frosch *et al.*, 2002). La littérature soulève également l'importance des faibles niveaux de pollution atmosphérique sur la santé. Même si l'Angleterre présente des niveaux de pollution inférieurs à la régulation, Janke *et al.* (2009) trouvent des effets des PM_{10} et de l'ozone sur les taux de mortalité. Aux Etats-Unis, les estimations de Currie & Neidell (2004) confirment que la pollution de l'air a un effet significativement positif sur la mortalité infantile, même à des niveaux de pollution relativement bas. Yu *et al.* (2001) fournissent une preuve additionnelle à l'existence d'effets sur la santé d'une exposition de long terme à un faible niveau de pollution à Hong Kong. Le projet E.R.P.U.R.S en France montre un effet positif du NO_2 et des PM_{10} sur le nombre d'admissions à l'hôpital, même à de faibles niveaux de concentrations (Campagna *et al.*, 2003). Pascal *et al.* (2009), en France, obtiennent des résultats similaires en considérant les taux de mortalité dans neuf villes françaises. Dans ce contexte, nous nous interrogeons sur les niveaux de pollution nécéssaire pour déclencher un effet sur la santé humaine. La réaction sanitaire à un effet de la pollution peut varier fortement selon la vulnérabilité des individus. Les enfants, les femmes enceintes, les personnes âgées, les asthmatiques, les insuffisants respiratoires et cardiaques et les fumeurs sont plus fragiles face à la pollution atmosphérique (Airparif, 2013). En matière de qualité de l'air, les valeurs limites, les seuils de recommandation, d'information, d'alerte existent et visent à éviter, prévenir ou réduire les effets nocifs sur la santé humaine ou sur l'environnement dans son ensemble. A titre d'exemple, le SO_2, responsable d'un certain nombre de problèmes respiratoires, est réglementé en Europe comme aux Etats-Unis à $350\mu g/m^3$ et à 75 parties par milliards en moyenne horaire, respectivement. Mais les seuils actuels permettent-ils de protéger efficacement la santé humaine?

Un deuxième angle d'attaque de la thèse concerne les différences d'exposition face à la pollution atmosphérique.

Des études récentes trouvent une association positive entre le niveau de revenu et la santé (Subramanian & Kawachi, 2006), (Gunasekara *et al.*, 2011). L'éducation est également considérée comme une composante cruciale du statut socio-économique (Grossman, 2000). Stringhini *et al.* (2012) suggèrent que les comportements sanitaires contribuent fortement aux différences dans les niveaux de santé parmi les catégories socioéconomiques. Parmi d'autres, Lindahl (2005) trouve un lien positif et causal entre le niveau de revenu et les taux de mortalité. Viel *et al.* (2011) soulignent aussi que les lieux avec de fortes proportions d'immigrants présentent une tendance à l'accumulation de sites polluants en France. En Allemagne, Schikowski *et al.* (2008) constatent l'existence de différences sociales dans la santé respiratoires et Bolte *et al.* (2010) décèlent des inégalités sociales de perception à l'exposition environnementale du logement. Pearce *et al.* (2010), en Nouvelle-Zélande, révèlent que même si la pollution industrielle est plus importante dans les zones riches, la pollution en général affecte plus les zones pauvres. Les inégalités environnementales sont devenues un thème fondamental qui guide le développement des politiques économiques en France (World Health Organization (WHO), 2012). Et les facteurs environnementaux influant sur la santé, comme l'exposition à la pollution, n'ont pas été étudiés en France dans une approche en termes d'inégalités sociales. Les populations défavorisées peuvent ainsi être soumises à des risques environnementaux particulièrement élevés. Dans ce contexte, nous pouvons nous demander si les effets locaux de la pollution sur la santé s'accompagnent d'inégalités sociales en France. Observe-t-on des inégalités sociales de santé liées aux effets de la pollution sur la santé en France? Ces inégalités sociales s'accumulent-elles au sein des départements ou des communes françaises?

Les différences observées d'exposition à la pollution peuvent être liées aux différences de préférences pour la qualité de l'air. La question des disparités environnementales nous amène à nous interroger sur la possibilité d'évaluer les préférences pures des agents en matière d'environnement. Qui s'intéresse le plus aux problèmes de pollution et quel est le rôle de la contrainte budgétaire des agents dans la réponse à un changement environnemental? Selon Baumol & Oates (1988), les pauvres et les riches accordent une importance différente à la protection de l'environnement. Cette idée tient tout d'abord son origine des différences de revenus qui existent entre les deux catégories et leurs impacts sur le consentement à payer (CAP) des agents. Elle dépend aussi du type de bien que l'on considère et de l'exposition à la pollution des agents. La plupart des évidences théoriques et empiriques nous montrent que les classes plus aisées se sentent plus concernées par les problèmes environnementaux que les plus pauvres. La sensibilité à l'environnement a d'ailleurs souvent été initié par des catégories aisées disposant d'un statut important au sein de la société (Scruggs, 1998). L'absence de langage environnementaliste dans les discussions ne peut cependant pas être considérée comme un manque d'engagement environnemental. Les recherches économiques et sociales suggèrent aussi que les préférences environnementales sont hétérogènes à l'intérieur même des catégories de revenus (Pearce *et al.*, 2010). Or, la plupart des études tendent à agréger les données individuelles et nous empêchent de mesurer la réelle sensibilité environnementale des agents. Par exemple, lors du calcul du CAP, la plupart des études se focalisent sur le consentement à payer moyen ou médian pour évaluer les données. Le bon indicateur du consentement à payer n'est pas son niveau absolu, mais la part du revenu qu'il représente. En d'autres termes, le fait qu'un ménage pauvre soit prêt à payer une somme moins importante qu'un ménage aisé pour jouir d'un bien public, signifie-t-il qu'il a une priorité moindre pour la protection de l'environnement ?

Nous observons également une littérature économique florissante sur l'hypothèse selon laquelle la pollution de l'air et les conditions environnementales auraient un impact direct sur la productivité du travail. Quels seraient dans ce contexte les gains économiques d'une amélioration de la qualité de l'air? Gillette (1977) trouvait déjà que la concentration d'ozone dans le sud de la Californie avait réduit les bénéfices journaliers de la récolte de citrons. En ce sens, une amélioration de la qualité de l'air peut avoir une influence positive sur le nombre d'heures travaillées en diminuant l'absentéisme au travail (Currie *et al.*, 2009). L'idée étant que la santé joue un rôle prépondérant sur l'efficacité au travail. Hanna & Oliva (2011) montrent récemment que la fermeture d'une raffinerie au Mexique, suite à une diminution de la pollution, a des effets positifs sur les salaires de la ville de Mexico. Graff Zivin & Neidell (2012) mettent en lumière l'impact de l'ozone sur la productivité au travail. Ils soutiennent l'idée que la protection environnementale peut aussi être vue comme un investissement en capital humain. Sa contribution à la productivité des entreprises et à la croissance économique devrait donc être incorporée au calcul des décideurs politiques.

1.3. Les données françaises disponibles

1.3.1. Le réseau français de mesure de la pollution atmosphérique

Le réseau français de surveillance de la qualité de l'air est composé de 38 associations, agréées pour la surveillance de la qualité de l'air (AASQA) réunies au sein de la fédération ATMO, certifiée par le Ministère de l'Ecologie et du Développement Durable (MEDD). Les données utilisées sont celles de l'Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME) et plus récemment de l'Institut National de l'Environnement Industriel et des RISques (INERIS) qui assure la coordination technique de la surveillance de la qualité de l'air au niveau national et gère la Base de Données Nationale de Qualité de l'Air. Les principaux polluants atmosphériques, dioxyde de soufre (SO_2), particules en suspension (PM), oxydes d'azote (NOx), ozone (O_3), monoxyde de carbone (CO), dioxyde d'azote (NO_2) sont soumis à leur contrôle car ils représentent un risque environnemental majeur pour la santé et leur nocivité a été clairement démontrée dans de nombreuses études. Faisant l'objet de réglementations, ce sont ces indicateurs de pollution atmosphérique que nous avons choisis pour construire notre base de données.

1.3.2. Les données de Météo France

Les variables météo proviennent de Météo France. Elles sont présentes dans la plupart des estimations pour contrôler les variations de la pollution dans l'air. La pollution de l'air ambiant peut varier énormément, les variations s'expliquent en partie par la variabilité des émissions atmosphériques mais surtout par la complexité des phénomènes météorologiques qui contrôlent la dispersion des polluants ou au contraire leur accumulation. Vent et pluie favorisent la dispersion des polluants. En revanche, une pression atmosphérique élevée, de fortes températures peuvent se traduire par une concentration progressive des polluants. Les variables météo que nous utilisons dans cette thèse (précipitations, humidité, température et vent) agissent comme des facteurs de confusion qui peuvent fortement influencer la concentration de la pollution dans l'air.

1.3.3. Le programme de médicalisation des systèmes d'information

Le programme de médicalisation des systèmes d'information (PMSI) en France donne accès à des informations administratives et médicales sur les admissions à l'hôpital en France. Depuis la loi du 31 juillet 1991 portant sur la réforme hospitalière, les établissements de santé publics et privés doivent procéder à l'analyse de leur activité médicale et transmettre aux services de l'état et à l'Assurance maladie "les informations relatives à leurs moyens de fonctionnement et à leur activité"(articles L. 6113-7 et L. 6113-8 du code de la santé publique). Il est stipulé qu'ils doivent "mettre en oeuvre des systèmes d'information qui tiennent compte notamment des pathologies et des modes de prise en charge". Tout séjour hospitalier en soins de courte durée − médecine, chirurgie, obstétrique et odontologie (MCO) d'un établissement de santé public ou privé doit donner lieu à la production d'un résumé de sortie standardisé (RSS). Les données d'admissions hospitalières pour les établissements publics et privés sur l'ensemble de la France pour l'ensemble des pathologies de 2006 à 2011 ont été obtenues auprès de l'agence technique de l'information sur l'hospitalisation (ATIH) à travers le programme de médicalisation des systèmes d'information (PMSI) et suite à une demande faite à la CNIL. Les données de santé utilisées sont strictement anonymes car la seule information disponible est le code postal du patient. Par ailleurs, le recul limité, la montée en charge progressive des services participant à ce système, et la qualité et le codage de certaines variables, peuvent entraîner des limites dans les analyses dont nous resterons conscients lors de l'étude économétrique.

1.3.4. Les données de prix de l'immobilier

Les notaires alimentent un fichier des actes de ventes immobiliers en France. La société PERVAL pour la province et l'Ile-de-France gère cette base de références immobilières. La base de données immobilières des notaires rassemble plus de quatre-vingt variables. Cette base référence l'ensemble des transactions immobilières françaises pour les appartements, maisons, locations et terrains. Nous utilisons dans cette thèse les données microéconomiques de l'ensemble des transactions immobilières effectuées dans la région Nord-Pas de Calais entre 2008 et 2011. Ces micro données détenues par la chambre des notaires sont extrêmement riches car elles contiennent non seulement un nombre important de variables sur les appartements et les maisons (volume de ventes enregistrées, le prix au m^2, le prix de vente, la surface habitable, la ventilation par nombre de pièces, etc...), mais aussi des informations socioéconomiques sur l'acquéreur. La catégorie socioprofessionnelle, la provenance et la nationalité définissent les caractéristiques des intervenants et facilitent également une lecture géographique des disparités sociales.

1.4. Les différentes méthodes empiriques utilisées

1.4.1. Les données de panel et les modèles à effets fixes

Cette thèse s'inscrit dans cette littérature récente sur la santé qui étudie les relations entre santé et pollution atmosphérique en insistant sur l'impact de la pollution de l'air sur les taux de mortalité en Angleterre (Kunzli *et al.*, 2000), (Janke *et al.*, 2009), sur la sensibilisation allergique des enfants dans les écoles primaires françaises (Annesi-Maesano *et al.*, 2007), sur l'asthme (Wilhelm *et al.*, 2009), ou les risques de cancer parmi les enfants scolarisés aux Etats-Unis (Chay *et al.*, 2003), (Morello-Frosch *et al.*, 2002). La grande majorité des études sur ce sujet utilise des séries temporelles ou des cohortes transversales (Janke *et al.*, 2009). Les séries temporelles exploitent les variations de court terme afin d'identifier les effets de la pollution sans pouvoir éliminer l'influence de facteurs de confusion comme le tabac, l'activité physique ou l'alimentation car ces facteurs ne changent pas sur le court terme. Les études de cohortes souffrent également d'un biais lié aux variables omises car les municipalités ou les zones géographiques comparées diffèrent certes, en termes de pollution, mais aussi en termes d'activités socioéconomiques, de politiques locales, de démographie, etc...

L'utilisation d'un modèle multivarié permet aux coefficients de régression d'être analysés de manière simultanée à différentes échelles spatiales. Nous utilisons un modèle à effets fixes individuels qui permet de capter l'essentiel des différences entre zones géographiques (département ou commune), non prises en compte par les facteurs de confusion. L'économétrie spatiale retient également toute notre attention. Nous utilisons la statistique globale de

Moran pour mesurer l'autocorrélation spatiale présente dans nos données puis nous estimons un modèle considérant l'existence d'hétérogénéité spatiale du territoire français. Par ailleurs, nous discutons le problème d'endogéneité fréquemment abordé dans la littérature économique environnementale et présent dans ce type de modèle.

1.4.2. L'expérience naturelle

Cette thèse est également reliée à la littérature récente sur la santé se préoccupant des problèmes de sélection non aléatoires et des réponses comportementales liées aux préférences pour la pollution. Les études économiques récentes utilisent des événements exogènes pour ls traiter. Chay & Greenstone (2003) analysent l'impact de la pollution de l'air sur la santé des enfants suite à une récession et concluent que la réduction de la concentration en TSP d'1% entraîne une diminution de 0.35 % de la mortalité infantile. Hanna & Oliva (2011) considèrent les effets sur les salaires de la fermeture d'une raffinerie au Mexique; elles observent qu'une augmentation d'1% de dioxyde de soufre engendre une diminution de 0.61 % d'heures travaillées. Moretti & Neidell (2011) utilisent les variations de pollution du port de Los Angeles pour étudier son impact sur la santé. Leurs estimations montrent que l'ozone à un coût annuel d'au moins quarante-quatre millions de dollars à Los Angeles en termes d'admissions à l'hôpital pour pathologies respiratoires. Currie *et al.* (2011) examinent l'impact du programme américain Superfund pour le nettoyage des sites pollués sur la santé des enfants et trouvent que le risque de maladies congénitales à côté des sites pollués, avant la mise en place du programme Superfund, était de 20 à 25% supérieur aux zones à plus de deux kilomètres des sites pollués. Greenstone & Gallagher (2008) analysent également les effets du programme Superfund sur le prix de l'immobilier. Ils observent que les évolutions des prix de l'immobilier, de l'offre de maisons, de la population, ne sont pas significativement liées au programme Superfund. Currie & Walker (2011) étudient l'impact de la diminution de la pollution suite à l'instauration du télépéage sur la santé des enfants. Celle-ci réduit le nombre de prématurés et le faible poids des enfants à la naissance dans les zones situées à moins de deux kilomètres par rapport aux zones situées entre deux et dix kilomètres. Récemment, Currie *et al.* (2013) analysent l'impact de la fermeture et de l'ouverture de 1600 sites industriels polluants sur le marché de l'immobilier et la santé. La valeur des maisons situées autour d'un mile d'un site industriel augmente d'environ 1,5% lorsque le site industriel ferme et diminue d'environ 1,5% lorsque le site industriel ouvre.

Les sources de pollution affectant la santé sont nombreuses et une des contributions principales de cette thèse est de mettre en lumière l'impact sanitaire d'une importante concentration de la pollution de l'air en utilisant des expériences naturelles. Grâce aux données de panel disponibles en France, un moyen d'isoler l'effet causal de la réduction de concentration en pollution de l'air liée à l'arrêt d'un site industriel est d'examiner les différences en termes d'impact entre les communes proches de ce site industriel et les communes plus éloignées de ce site. Il est en effet difficile de comparer l'état de santé à proximité d'un site polluant avec l'état de santé sur un autre site pour mesurer l'impact de la pollution. En effet, les populations vivant à proximité d'un site polluant peuvent avoir des comportements à risque qui pourraient expliquer leur mauvais état de santé. L'arrêt d'une activité polluante et l'utilisation d'un modèle de différence en différence nous permet de bien mesurer l'effet de la pollution sur la population habituellement exposée à la pollution provenant d'un site polluant par rapport aux populations vivant loin du site et donc non exposées à celui-ci.

Néanmoins, l'utilisation de la méthode de différence en différence repose sur de nombreuses hypothèses qu'il est crucial de ne pas négliger. Puisque l'estimateur utilisé est une variable d'interaction entre la période de fermeture et la commune du site pollué, l'hypothèse d'identification s'appuie sur le fait que les zones proches et éloignées du site pollué ont des tendances similaires en termes de pollution et de santé au sein de la même région. Cela peut ne pas être le cas si, par exemple, les zones à la périphérie du site pollué se développent à un rythme plus rapide que les zones plus proches de la raffinerie ou abritent moins d'industries polluantes. Cependant, la structure de panel nous permet des contrôles supplémentaires qui purgent de tous biais liés à une évolution différente. Par ailleurs, dans ces modèles, l'hypothèse d'exclusion serait violée si d'autres événements avaient eu lieu en même temps que la fermeture de la raffinerie. Il est donc nécessaire qu'aucun événement économique ou politique ne coïncide avec le moment exact de l'arrêt des activités et l'emplacement du site. L'hypothèse d'exclusion peut toutefois encore être violée si la fermeture du site elle-même a un impact sur les variables de santé indépendamment de la pollution. Par exemple, les travailleurs au chômage peuvent avoir plus de temps et ainsi influencer le nombre d'admissions à l'hôpital. En outre, des migrations dans les zones entourant le site peuvent s'être produites après la fermeture et biaiser les résultats.

1.5. Une thèse en trois chapitres

Le premier objectif de cette thèse est de rassembler deux pans de la littérature en se focalisant sur les liens existant entre pollution et risques sanitaires et sur les inégalités sociales d'exposition aux risques environnementaux. Par le biais d'études économétriques, nous étudions d'une part les taux de mortalité associés à la pollution atmosphérique selon le statut socioéconomique de la population des départements français et, d'autre part, les liens entre pollution atmosphérique et santé des enfants à la naissance au sein des communes françaises en utilisant une expérience naturelle. Par ailleurs, la dernière partie de cette thèse s'attache à mettre en regard la méthode des prix hédoniques, évaluation des références reposant sur une perception subjective des effets de la pollution, avec la fonction de dommage, méthode de valorisation objective des coûts de la santé.

1.5.1. Pollution atmosphérique, disparités environnementales et mortalité

Le chapitre 1 met en exergue des inégalités environnementales, à l'échelle nationale, en France métropolitaine, et avec pour unité spatiale le département. A travers une analyse économétrique de données de panel pour la période 2000-2004, les taux de mortalité sont étudiés en relation aux immissions de polluants dans l'air ambiant et au statut socioéconomique. Les niveaux de concentration de CO, SO_2, NO_2, NO, O_3 et PM_{10} sont estimés par interpolation spatiale à partir d'un réseau de stations de mesure de la pollution. Par le biais d'un modèle multivarié, nous étudions les relations entre les taux de mortalité et ses principaux déterminants pour ensuite faire le lien avec la pollution atmosphérique mesurée pour chaque département. Nous trouvons un impact positif et très significatif sur le taux de mortalité non-accidentel, plus particulièrement lorsque l'on se situe à de forts niveaux de concentration de pollution avec un risque relativement plus important pour les femmes. Par ailleurs, le dioxyde d'azote (NO_2) agit comme une variable modératrice de l'influence du taux de chômage sur les taux de mortalité.

Cet article contribue à enrichir la littérature actuelle sous plusieurs angles. La plupart des études empiriques internationales se focalisent, d'une part, sur les liens existants entre pollution et risques sanitaires ou, d'autre part, sur les inégalités sociales d'exposition aux risques environnementaux. L'objectif de cette étude consiste à rassembler ces deux courants de la littérature afin d'aboutir à des différences de risques sanitaires associés aux externalités environnementales selon le statut socioéconomique de la population des départements français. Les facteurs environnementaux influant sur la santé, comme l'exposition à la pollution, n'ont pas été étudiés en France à l'échelle nationale dans un contexte d'inégalités sociales. De plus, contrairement à la plupart des études épidémiologiques précédentes, nous utilisons un modèle à effets fixes individuels et un modèle prenant en considération l'autocorrélation spatiale ce qui permet de rendre compte de l'essentiel des facteurs inobservés et de capter l'hétérogénéité spatiale qui caractérise le territoire français.

Cette étude s'intéresse principalement à la pollution de fond et les niveaux de concentration annuels départementaux sont estimés par interpolation spatiale à partir des données du réseau de stations de mesure de la pollution. Le poids attribué à chaque station correspond à l'inverse de la distance entre la préfecture et la station de mesure. Cette mesure de concentration de pollution pour chaque département nous permet de faire le lien avec les taux de mortalité standardisés à l'échelle départementale .

La stratégie empirique de cet article consiste à étudier les relations entre les principaux déterminants de la mortalité pour ensuite faire le lien avec la qualité de l'air ambiant de chaque département français. Nous estimons dans un premier temps un modèle standard de la mortalité non accidentelle avec ses principaux déterminants (comportements à risques, niveau de vie, niveau d'industrialisation, météorologie, accès aux soins). Puis, nous ajoutons à la régression les variables de pollution atmosphérique dans un modèle à simple et à plusieurs polluants afin d'étudier leur impact sur les taux de mortalité. Les variables explicatives précédentes font office de variables de contrôle. Nous distinguons également un modèle à effet genre. Les niveaux de concentration de pollution sont aussi divisés par la médiane en deux catégories de risque afin de différencier l'impact selon l'intensité d'immission. Finalement, l'inclusion d'une variable d'interaction nous amène à nous interroger sur les liens entre le statut socioéconomique et lŠeffet de la qualité de lŠair.

Grâce à l'utilisation d'outils statistiques appropriés, nos résultats supportent fortement l'idée que le dioxyde d'azote a un impact positif et significatif à 1% sur le taux de mortalité, plus particulièrement lorsque l'on se situe à de relativement forts niveaux de concentration de pollution. Toutefois, les immissions de polluants étudiées se situent en dessous des seuils de pollution fixés par les autorités publiques, confortant l'idée que la pollution est fluctuante et que même des niveaux faibles de concentration peuvent être actifs. Nous mettons également en évidence des risques sanitaires plus importants chez les femmes pouvant être justifiés par des différences dans les types d'activités. Par ailleurs, le NO_2 agit comme une variable modératrice de l'influence du taux de chômage sur les taux de mortalité.

Ces résultats ne confirment pas seulement l'existence de relations de long terme entre la pollution de fond et la mortalité mais soulèvent également des questions importantes à propos de la justice environnementale en France.

1.5.2. L'impact de la production d'énergie et les externalités de santé

Ce deuxième chapitre s'intéresse aux externalités de santé liées à la production d 'énergie. Le but est d'explorer les effets des externalités environnementales sur la santé et de souligner leurs conséquences sur la santé des enfants à la naissance (poids, âge gestationnel) grâce à une approche empirique géographique fine. A travers une analyse économétrique de données de panel, les bases de résumés anonymes du programme de médicalisation des systèmes d'information seront étudiées en relation avec les immissions de polluants dans l'air.

Nous examinons l'impact du dioxyde de soufre (SO_2) en France sur la santé des enfants à l'échelle de la commune. Nous utilisons les grèves d'octobre 2010 en France dans les raffineries françaises comme expérience naturelle. Cette étude fait partie d'une littérature récente sur la santé s'intéressant aux biais de sélection et d'autosélection liés aux préférences pour la pollution. La première contribution importante de ce papier est de mettre en lumière l'impact sanitaire d'une importante concentration en dioxyde de soufre qui résulte de la production d'énergie. D'autre part, ce choc exogène sert à étudier l'impact sanitaire d'une importante concentration en SO_2.

La demande énergétique internationale continue à augmenter et la source d'énergie la plus utilisée reste le pétrole, particulièrement pour le transport. Pour répondre à cette demande, de nombreux traitements transforment le pétrole brut en produits à haute valeur commerciale. La production d'énergie provenant des raffineries pose néanmoins un risque sanitaire. Les émissions incluent un certain nombre de polluants avec des effets notoires sur la santé, plus particulièrement le SO_2. Le processus de raffinage rejette en effet du SO_2 dans l'air. En France, 20 % des émissions en SO_2 proviennent de la production d'énergie. Les rejets en dioxyde de soufre sont dus pour une large part au secteur industriel.

Ce deuxième chapitre se focalise donc sur le SO_2, l'un des polluants majeurs émis par les raffineries de pétrole et le principal polluant du secteur industriel. Le SO_2 est responsable d'un certain nombre de problèmes respiratoires, c'est pourquoi ce polluant est réglementé sous la norme américaine et française de propreté de l'air.

Il existe onze raffineries en France et celles-ci traitent plus de quatre-vingt-neuf millions de tonnes de pétrole chaque année. En octobre 2010, les salariés des raffineries françaises se sont mis en grève suite à un mouvement social contre la réforme des retraites. La grève a débuté mi-septembre 2010 pour durer un mois et demi. Alors que certaines raffineries ont réduit leur activité de raffinage à un minimum pendant cette période, d'autres ont arrêté complètement leur travail pendant dix-sept jours, fin octobre. Nous utilisons ce choc exogène pour faire le lien entre pollution atmosphérique et santé des enfants à la naissance.

Notre objectif est d'analyser l'impact d'une réduction de la pollution de l'air sur la santé pour les communes qui ont bénéficié de cette réduction de pollution suite aux grèves. Le meilleur moyen d'isoler l'effet causal de la réduction de concentration en dioxyde de soufre liée aux grèves est d'examiner les différences en termes d'impact entre les communes avec une raffinerie et les communes sans raffinerie. Le SO_2 , même à des niveaux inférieurs aux seuils de régulation dans la plupart des pays du monde, a des effets statistiquement significatifs et positifs sur l'âge gestationnel et le poids des enfants à la naissance. Cette étude fournit un support solide à l'idée que les populations proches des raffineries sont plus exposées à la pollution et plus gravement affectées que les populations plus éloignées des raffineries.

1.5.3. Le prix de la pollution et la santé

Le troisième chapitre met en regard le consentement à payer des agents pour une amélioration de la qualité environnementale et les bénéfices économiques objectifs en termes de santé d'une amélioration de la qualité environnementale. Il s'attache à mettre en regard la méthode des prix hédoniques avec la fonction de dommage. La méthode "fonction de dommage" tient compte de la pollution observée (bénéfice environnemental objectif) tandis que la méthode des prix hédoniques tient compte de bénéfices environnementaux subjectifs. A partir de ce critère, nous analysons les différences de diagnostic entre les deux méthodes.

Ce travail empirique s'intéresse aux externalités de santé liées à la production d'énergie et à l'évaluation de leur importance pour les agents économiques. L'objectif est d'explorer les conséquences des externalités de santé sur les prix de l'immobilier grâce à une approche empirique géographique fine. Pour ce faire, nous examinons l'impact du dioxyde de soufre (SO_2) dans la région Nord-Pas de Calais en France sur les admissions à l'hôpital pour pathologies respiratoires et les prix de l'immobilier à l'échelle de la commune pour la période de 2007 à 2011. A travers une analyse économétrique de données de panel, les bases de résumés anonymes du programme de médicalisation des systèmes d'information (PMSI) et la base PERVAL de la chambre des notaires sont étudiées en relation avec les immissions de polluants dans l'air et les facteurs locaux des communes. L'arrêt des activités de la raffinerie des Flandres dans le nord de la France, en septembre 2009, est utilisé comme l'expérience naturelle d'une amélioration de long terme de la qualité environnementale. Notre objectif est d'analyser l'impact d'une réduction de la pollution de l'air sur la santé pour la commune de Mardyck associée à Dunkerque, qui a bénéficié d'une réduction de pollution suite à l'arrêt de la raffinerie des Flandres en septembre 2009. En effet, encore une fois, le meilleur moyen d'isoler l'effet causal de la réduction de concentration en dioxyde de soufre liée à l'arrêt des activités de la raffinerie est d'examiner les différences en termes d'impact entre la commune d'implantation de la raffinerie et les autres communes. Tout d'abord, l'arrêt des activités de la raffinerie influence les mesures locales de pollution atmosphérique. L'utilisation d'un choc exogène répond aux problèmes de sélection non aléatoire qui peuvent avoir biaisé les résultats des études précédentes.

Cette étude montre, d'une part, que l'arrêt de l'activité de raffinage, suivi en 2010 par la fermeture définitive de la raffinerie, diminue la concentration en SO_2 et a des effets significatifs sur la réduction de la sévérité des pathologies respiratoires. Ce résultat est particulièrement significatif pour les populations à risque comme les jeunes enfants et les personnes âgées de plus de soixante-dix ans. La santé d'une partie de la population active semble aussi positivement influencée par cette diminution de pollution suggérant un lien avec la productivité du travail.

D'autre part, ce choc exogène nous permet d'examiner si ce bénéfice économique, en termes de santé, suite à une amélioration de la qualité de l'air, se reflète bien dans l'évolution des prix de l'immobilier comme le prédit la théorie sur l'approche hédonique. Les résultats sur certains segments du marché ne sont pas ceux attendus. Nous observons tout d'abord, une absence de résultats significatifs quand la méthode des prix hédoniques est appliquée sur l'ensemble du marché immobilier. Le prix des maisons individuelles et des appartements onéreux augmentent significativement suite à la fermeture de la raffinerie dans la commune de Dunkerque. Etant donné les effets objectifs et positifs sur la santé, nous obtenons des résultats contraires pour le prix des appartements peu chers. Les différences en termes de subjectivité et d'objectivité ne permettent pas d'expliquer la différence de comportements entre les deux segments. L'évaluation de l'effet environnemental est moins importante en termes relatifs pour les propriétaires/acheteurs d'appartements que celle de l'effet économique. La contrainte budgétaire semble plus serrée pour les propriétaires d'appartement, qui auraient des dépenses contraintes plus importantes par rapport à leur budget que les autres.

Chapter 2

An Econometric Analysis of Atmospheric Pollution, Environmental Disparities and Mortality Rates

This paper presents the first study of environmental inequality related to health in France on the national scale. Through econometric analysis based on panel data from 2000 to 2004, at the level of France's departments (administrative areas similar in size to counties), I investigate the total mortality rate in relation to socioeconomic status and air pollution. The concentration level of NO_2, O_3 and PM_{10} are estimated by spatial interpolation from local observations made by a network of monitoring stations. I found a positive and significant relationship between NO_2 levels and the mortality rate, at mean levels below the current standard, with a greater relative risk for women. Moreover I observed disparities in health related to income among French departments. These results not only confirm the existence of a relationship between current air pollution levels and mortality but also raise questions about environmental policy implications in France.

2.1. Introduction

The prevalence of many pollutants is declining throughout the industrialized world. However, exposure to air pollution, even at the levels commonly reached nowadays in European countries, still leads to adverse health effects. In this context, there has been increasing global concern over the public health impacts attributed to environmental pollution.

Multilevel modelling has been previously used to assess the negative correlation between pollution exposure and socioeconomics status, such as unemployment, education, and the occupational class in Canada (Premji *et al.*, 2007), ethnic group, and income in England (McLeod *et al.*, 2000) and in the US (Grineski *et al.*, 2007), (Morello-Frosch *et al.*, 2002) where the concept of environmental justice has been the object of increasing attention. Viel *et al.* (2011) emphasize that towns with high proportions of immigrants tend to host more hazardous sites even when controlled for population size, income, degree of industrialization of the town, and region. In Germany, Schikowski *et al.* (2008) show the existence of social differences in respiratory health in the female population and Bolte *et al.* (2010) argue social inequality exists in terms of perceived environmental exposure in relation to housing conditions. Pearce *et al.* (2010) for New Zealand point out that industrial pollution is greater in wealthy areas, whereas overall pollution affects poorer zones more. Mohai *et al.* (2009) indicate that black respondents, those making less than $40,000 per year and those without a high school diploma were significantly more likely to live near an industrial facility. Pastor *et al.* (2001) also found a strong correlation between periods of greatest community demographic change and the introduction of noxious land uses that they called ethnic churning. Social factors such as employment status or access to social capital can play a major role in determining the response to environmental aggression. Harper *et al.* (2009) acknowledge that local sites for externalizing environmental harm are created when some landscapes and social groups are perceived as what they call "beyond the pale" of environmental regulation, public participation and civil rights. Finally, Finkelstein *et al.* (2003) point out that mean pollutant levels tend to be higher in lower income neighborhoods in Ontario Finkelstein *et al.* (2003).

Moreover, multiple models also estimate the relationship between health and pollution, showing the impact of outdoor air pollution on the mortality rate in England (Kunzli *et al.*, 2000), on the allergy sensitivity of primary-school children in France (Annesi-Maesano *et al.*, 2007), on asthma (Wilhelm *et al.*, 2009), or on cancer risks among school children in the US (Chay *et al.*, 2003), (Morello-Frosch *et al.*, 2002). The literature also claims low levels of pollution can be harmful. Although the UK has low current air quality limit values for airborne pollutants by historical standards, Janke *et al.* (2009) find that higher levels of PM10 and ozone are associated with higher mortality rates. For the US, Currie & Neidell (2004)'s estimates confirm that air pollution has a significant effect on infant mortality, even at relatively low levels of pollution. Their estimates suggest that the reductions in CO and PM_{10} that occurred during the 1990s saved more than 1 000 infant lives in California. The study from Yu *et al.* (2001) provides additional evidence for the adverse effects of long-term exposure to relatively low-level air pollution in Hong Kong. The E.R.P.U.R.S project in France shows that NO_2 and PM_{10} have a negative impact on health, even at low air concentrations, considering hospitalization numbers as the explicative variable (Campagna *et al.*, 2003). Pascal *et al.* (2009) in France obtain similar results, considering different mortality rates in nine polluted cities.

In addition, the literature is far from being silent about the relationship between health and socioeconomic status (SES). A number of SES measures have been proposed, including income, wealth, education, labor force status, and race/ethnicity. For instance, some recent papers provide evidence of a positive association between income and health (Subramanian & Kawachi, 2006), (Gunasekara *et al.*, 2011). Apart from income, education has also been considered a crucial component of SES affecting health (Grossman, 2000). In France, Cambois & Jusot (2011) study the link between lifelong adverse experiences, health and SES. Lifelong adverse experiences are related to poor self-perceived health, diseases and activity limitations, even controlling for SES. Results from Stringhini *et al.* (2012) suggest that the social patterns of unhealthy behavior differs between countries. They stress health behavior is likely to be a major contributor of socioeconomic differences in health. Among others, Lindahl (2005) focus on mortality rates and find a positive causal relationship between income and health measures.

Moreover, I observe a growing epidemiologic literature about the effects of air pollution on health by gender. The most recent gender analysis from Clougherty (2010) shows that most studies for adults report stronger effects among women, particularly when using residential exposure assessment. The smaller size of the trachea has been put forward as the reason which makes women more sensitive to particulates in the air (Marr, 2010). However, it remains unclear whether the observed difference is a result of gender-linked biological differences or gender differences in activity patterns.

The analysis offers several contributions to the existing literature. Most international empirical economic studies estimate either the relationship between health and pollution or the correlation between pollution exposure and socioeconomic status. I aim to bring together both aspects of the literature to assess the impact of air pollution on health according to socioeconomic status. Few European studies investigated the modification effect of socio-economic factors on the association between air pollution and health and much is yet to be understood (Deguen & Zmirou, 2012). European policy-makers have in fact only recently acknowledged the notions of environmental justice and environmental inequalities, which have been part of the US policy arsenal for almost two decades (Laurent, 2011). Results from Morello-Frosch *et al.* (2002) indicate that children of color, namely, Latinos and African-Americans, bear the highest burden of estimated cancer and non-cancer health risks associated with exposure to ambient toxic air, while they are at school in California. Finkelstein *et al.* (2003) also point out that mean pollutant levels tend to be higher in lower income neighbourhoods in Ontario and both income and pollutant levels are associated with mortality differences. To my knowledge, environmental factors affecting health, such as exposure to atmospheric air pollution have not been yet studied in France on a national scale in the context of social inequalities. Laurent *et al.* (2007) emphasize the importance of continuing to investigate this topic due to the tendency for greater effects to be observed among the more deprived. Whereas the French literature only looks at high levels of pollution, I study ambient air pollution: the dataset presents low levels of pollution concentration, below the actual threshold fixed by the public authorities, and at which health is said to be harmed (Pascal *et al.*, 2009). Instead of looking at one geographical area, I examine recent relationships between pollution and health for the entire country using a panel dataset. I also account for unobserved confounding factors using fixed-effect clustering at the regional level, in order to avoid potential omitted variable bias. Most of the studies on this topic use times series data or a cross sectional cohort (Janke *et al.*, 2009). Times series exploit short-term variations to identify pollutant effects. This eliminates the effects of lifestyle factors such as smoking, exercise and diet, because these factors do not change in the short run. The cohort studies may also suffer from omitted variables bias, as the cities or zip codes which are compared may differ from each other in important ways other than just their levels of pollution. Some recent studies use an exogeneous event to cope with omitted variables biais (Chay & Greenstone, 2003) (Moretti & Neidell, 2011), (Currie & Walker, 2011). For instance, Chay & Greenstone (2003) use a sudden recession as an instrument to identify the effect of a medium-term reduction of pollution on infant mortality. Finally, I use a model which takes into account spatial autocorrelation that may have bias results from previous studies using panel data with cross-sectional correlation.

This paper investigates the relationship between ambient air pollutant concentrations, social class, and population mortality at the departmental level.[1] It is part of

1. in France, a "Département" corresponds to a local authority below the regional level, and is similar in size to a county

new research on environmental justice, and provides an overview of the distribution of environmental risks. To identify the social distribution of air pollution, the study compared the social characteristics (income, unemployment) and the concentration of air pollution among French local authorities for different levels of poverty. In this context, we first may ask whether poor areas are also the ones with low socioeconomic levels. Poor people may be more likely to live in areas where pollution may be higher, next to industrial districts for example. Due to budget constraints, unemployed people are also less likely to move from one area to another to avoid pollution. Secondly, the study also asks whether a change in pollution benefits health more in areas higher socioeconomic groups than areas with persons on low income levels. When it comes to poorer local authorities, is the health effect of an increase in air pollution twice as large? The main purpose is to figure out if inequalities tend to accumulate in French local authorities.

I find a positive and significant relationship between NO and the mortality rate, at mean levels below the current standard, with a greater relative risk for women. I show that higher income levels of French departments are associated with lower mortality levels. However, health disparities appear to be more related to socioeconomic factors than differences in sensitivity to pollution.

2.2. Medical perspective

The L.A.U.R.E (Law on Air and Rational Use of Energy) and the different European directives give priority to monitoring common air pollutants with a direct effect on health, such as Nitrogen Dioxide (NO_2), Nitrogen Oxides (NO), Ozone (O_3) and Particles (PM_{10}). As a result, I consider these pollutants in this paper. The contamination of the atmosphere by pollutants at the local and regional level is the result of three processes: emission, transmission, and air pollution concentration. Pollutants are first released at source with gases and particles which are emitted into the air. The pollutants emitted are then dispersed, or sometimes they can be chemically transformed in the atmosphere, creating new, secondary pollutants. Having combined with air and being diluted, they create a concentration of toxic levels of chemicals in the air, and these atmospheric pollutants are finally inhaled by humans, animals and plants.

First, Particulate Matter (PM) is made up of a number of components, including acids, organic chemicals, metals, and soil or dust particles. The size of particles is directly linked to their effect on health: PM_{10} (aerodynamic diameter less than 10 μm); $PM_{2.5}$ (aerodynamic diameter less than 2,5 μm) are the particles that generally pass through the throat and nose and enter the lungs. The $PM_{2.5}$ particles are the most dangerous. The effects of PM on health occur at levels of exposure currently being experienced by most urban and rural populations in both developed and developing countries. Chronic exposure to particles contributes to the risk of developing cardiovascular and respiratory diseases, as well as of lung cancer (World Health Organization (WHO), 2011). Once inhaled, these particles can affect the heart and lungs and cause serious health effects. Not only have many

European projects found a link between particles and mortality or morbidity (Peng *et al.*, 2004) (Katsouyanni *et al.*, 1995), but so have recent epidemiologic studies (Schikowski *et al.*, 2008), (Janke *et al.*, 2009), (Annesi-Maesano *et al.*, 2007).

Nitrogen Oxides (NOx) are the main indicator of transportation and stationary combustion sources, such as contamination by electricity utilities and industrial boilers. [2] NOx form when fuels are burned at high temperatures and include various Nitrogen compounds such as Nitrogen Dioxide (NO_2) and Nitric Oxide (NO). They play a crucial role in atmospheric reactions by creating harmful particulate matter, ground-level Ozone, acid rain, and eutrophication of coastal waters. NO_2 is produced by chemical transformation with NO and Ozone ($NO + O_3 = NO_2 + O_2$). Not only particle filters but also the rise of Ozone in the atmosphere increases NO_2 emissions (AFSSET). As a consequence, NOx is a powerful oxidizing gas, linked to a number of adverse effects on the respiratory system (Environmental Protection Agency (EPA), 2011).

Ozone (O_3) is an example of a secondary pollutant as it is formed when Hydrocarbons (HC) and Nitrogen Oxides (NOx) combine in the presence of sunlight. And excessive Ozone in the air can have a marked effect on human health. It can cause breathing problems, trigger asthma, reduce lung function and cause lung diseases

2. The spatial distribution of (NO_2) is generally not homogeneous within individual metropolitan areas. The primary reason for the observed heterogeneity in concentrations across an urban area is the substantially higher concentrations of NO_2 nearby sources, such as roads [Electric Power Research Institute 2009]

(World Health Organization (WHO), 2006). Breathing ozone can trigger a variety of health problems including chest pain, coughing, throat irritation, and congestion (Environmental Protection Agency (EPA), 2011). Recent epidemiologic studies emphasize the relationship between Ozone and the mortality rate (Janke *et al.*, 2009) and asthma exacerbation (Currie & Neidell, 2004), (Laurent *et al.*, 2007), (Wilhelm *et al.*, 2009).

2.3. Presentation of the dataset

I use data on the concentration of pollutants and mortality rates available at a local level for the whole of France.

Detailed data on atmospheric pollution come from the information system of the air quality measure (BDQA) used by the French Environment and Energy Management Agency (ADEME) and more recently from the National Institute of Industrial Environment and Risks (INERIS). They gather information coming from the 38 associations (AASQA), within the ATMO federation which monitors air quality. A large number of monitoring stations make up the federation. The French nomenclature identifies seven classes of stations, consistent with the various classifications defined at the European level: roadside, urban, industrial, near city background, national rural, regional rural, specific observations numbering 84, 286, 119, 138, 10, 62, and 13 respectively. Most of the monitoring stations are placed where the population is significant, apart from national rural monitoring stations.

The measure taken into consideration in the study is the annual mean concentration for pollutants within a calendar year (1st January to 31st December), calculated by each AASQA for each sensor, and measured in micrograms per cubic meter of air. In principle, the more disaggregated data is, the more desirable it is to cope with ecological inferences, but the health authority estimates are based on surveys with relatively small samples and are therefore less reliable. However, aggregate data may offer valuable clues about individual behavior. I divide my dataset into subsamples, in an attempt to deal with the problems of confounding and aggregation bias. This annual mean is calculated by the ASQAA from the hourly mean for each monitoring station. This unit of concentration is mostly used to monitor outdoor air quality. Air pollutant concentrations do not necessarily produce accurate predictions of exposure levels. People may be resident in one area, but work in another. Nevertheless, the geographical level used in this article reduces the bias related to population mobility. The department surface represents an average of 570 000 hectares and I know from French National Institute for Statistics and Economic Studies (INSEE) data that the average distance between the place of residence and the place of work is nearly 20km, so the accuracy of the exposure levels seems reasonable.

For spatial interpolation between monitoring stations, I use a geostatistical method that takes into account spatial dependence. This method does not necessarly reduce the amount of measurement error in the variable. The extent of measurement error is going to be greater for those departments with few monitoring stations where the population is more dispersed or lower. Lower or higher levels of the dependent variable within departments also induce measurement error. For example, more rural areas tend to be more agriculturally based and this may have an impact on mortality rates. Nevertheless, measurement error, even if not systematic, can induce attenuation bias.

Following Currie & Neidell (2004), I assign annual pollutant concentrations to the 95 French departments, using an inverse-distance weighted average of pollution measurements from monitors within the same department. Using the geographical coordinates of the census blocks with the highest population of a local authority, I calculate the distance between the census blocks with the highest population and all monitoring stations as explained in Appendix 1. This distance corresponds to the weight attributed to a monitoring station.

In order to assess the accuracy of our measure, I compare the actual level of pollution at each monitor location with the level of pollution that I would assign using the method previously described. The correlations between the actual and predicted levels of pollution are quite high for O_3, NO_2 and PM_{10} (0.6, 0.85 and 0.7 respectively) suggesting that the measure is quite accurate.

The top panel of Table 2.1 presents descriptive statistics for pollution data. NO, NO_2 and PM_{10} are positively correlated with correlation coefficients between 0.5 and 0.8, as can be seen in table 2.2. They are negatively correlated with O_3, which may

Table 2.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Pollutant variables					
NO_2 ($\mu g/m^3$)	31.714	11.375	12	74.046	220
O ($\mu g/m^3$)	53.029	15.39	30.601	99.653	220
PM_{10}($\mu g/m^3$)	21.461	5.508	8.818	57.384	220
Mortality rates					
Non-incidental mortality rate (per 100,000)	819.759	64.615	620	1000	220
Non-incidental Ms (per 100,000)	626.536	47.623	499	756	220
Non-incidental Mr (per 100,000)	1092.595	96.551	792	1393	220
Socioeconomics variables					
Income (%)	15303.625	2703.677	11011.659	27079.313	220
Un (%)	8.436	2.033	4.575	14.625	220
Education (%)	15.678	5.045	10.241	37.481	220
Poverty gap	0.5	0.501	0	1	220
Weather variables					
Sun (hours)	1979.783	362.492	1367.4	2962.3	205
Pr (mm)	2.211	0.571	0.855	3.865	214
Wind (km/hour)	100.212	13.056	68.400	147.6	208
Frost (days)	41.61	22.186	4	114	213
Demographic variables					
Sm (%)	1218.513	274.182	483.5	2298.7	220
Industry (%)	16.175	4.778	5.93	24.476	220
PPHB (%)	136.174	37.84	71.696	268.82	220
Alcohol (%)	262.614	223.659	60	1594	220
Atmo Index					
Atmo index 8 to 10	3.524	5.532	0	28	220

be due to the fact that Ozone is rapidly destroyed to form NO_2. Within cities, the correlation between both NO_2 and NO is high (0.85), so that I choose to keep NO_2 as an explanatory variable and drop NO_2 to prevent autocorrelation. Moreover, I do not include observations for SO_2 and CO, as few monitoring stations measure these pollutants.

Table 2.2: Pairwise correlation coefficients with significance level

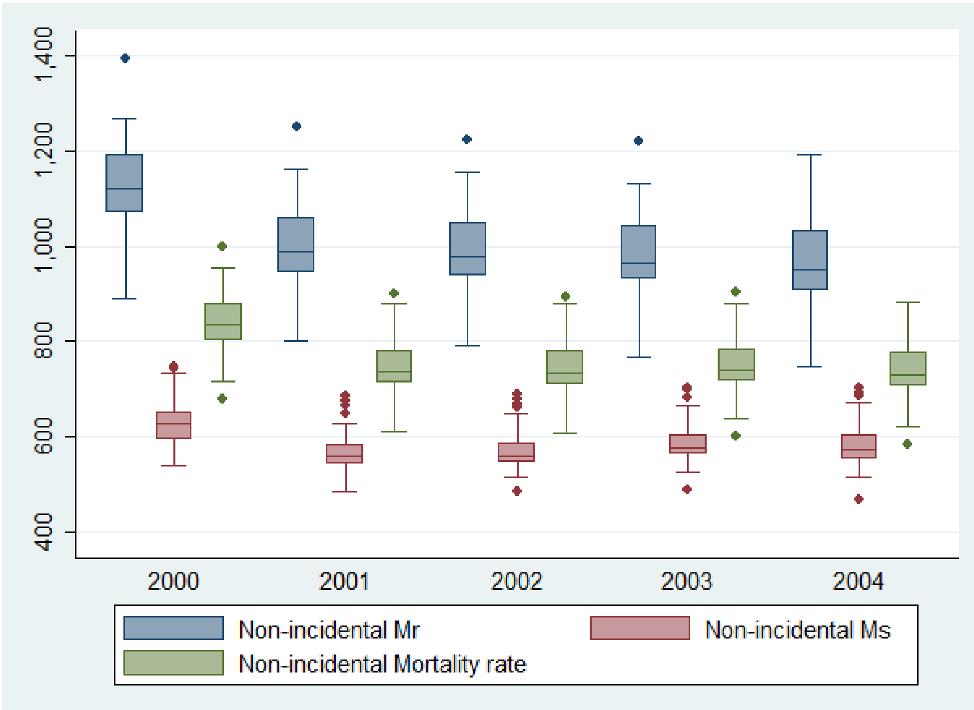
	NO_2	PM_{10}	O	NO	Atmo index(8-10)	Temperature
NO_2	1.0000					
PM_{10}	0.6086	1.0000				
	(0.0000)					
O	-0.3238	0.0301	1.0000			
	(0.0000)	(0.6566)				
NO	0.9644	0.5921	-0.3265	1.0000		
	(0.0000)	(0.0000)	(0.0000)			
Atmo index(8-10)	0.2663	0.5121	0.3833	0.1473	1.0000	
	(0.0034)	(0.0000)	(0.0000)	(0.1099)		
Temperature	0.1818	0.3079	0.4762	0.1447	0.2593	1.0000
	(0.0077)	(0.0000)	(0.0000)	(0.0344)	(0.0044)	

Other pollutants are also likely to be associated with differences in mortality, but data were unavailable to perform intra-urban interpolations for these pollutants. Note that the local authorities with missing air pollution measures are all less populated areas. It is important to stress that air pollutant concentrations used to be below the limit value fixed by European and national institutions, above which health may be harmed. In France, the threshold for NO_2, fixed by the European Act 2002-13 relating to air quality, is $200\mu g/m^3$ over 24 hours, to protect human health. For long term exposure (over one year) the regulated level is $40\mu g/m^3$ according to the WHO. It should be noted that the annual mean for NO_2 is below the regulated level. The annual mean is the measure I mainly use in my estimations. It corresponds to 'very good' air quality according to the ATMO index presented in table 2.4.

The second panel of Table 2.1 presents non-incidental mortality rates. I consider a period of 5 years (2000-2004). The year corresponds to the mid-years of the triennial period used. A moving average makes it possible to "smooth" a series of values expressed according to time. It is used to eliminate the least significant fluctuations. Mortality rate is a moving average of order 3. Data on mortality are available from 1980 to 2004 whereas data on pollution only exist from 1985 to 2005 with very few values before 2000. A large range of pollutants are responsible for outdoor air pollution, so that it is difficult to assign them to a specific health effect. As I wish to extract the only effect of pollution on mortality, I work on the non-incidental rate. This paper does not include specific causes of mortality, due to the weak variability of these data in France for the 2000-2004 period, which does not allow any estimations. The non-incidental rate does not take into account external causes of mortality as classified by the international classification of diseases (CIM10). For instance, transport releases $N0_2$ such that a high level of pollution may be observed next to roads where road accidents occur and could bias estimates if I was considering an overall mortality rate. The data on health come from the National Federation of Regional Health Observatories (ORS). I use age-standardized rates to control for different age structures across departments.[3] The standard deviation is quite high, showing that the data are spread out over a large range of values. The degree of dispersion (spread) and skewness in the data are presented graphically in Figure 2.1 .

3. This age-standardized rate is calculated as follows: $\sum_{i=1}^{19} PiTi$. Pi represents the share of age group for the population of reference and Ti represents the specific rate of mortality observed within a department for the age group i.

Figure 2.1: The yearly distribution of non-incidental mortality rates for the 2000-2004 period, in all departments.



The third panel presents the socioeconomic variables: income, education, poverty gap, poverty rate, standard of living and unemployment base, on the 2007 census by INSEE and the French Ministry of Labour (DARES). Definitions of the variables are given in the Table 1.5.3. It should be noted that data about ethnicity or race do not exist in France. The French Institute of Statistics does not collect data about language, religion, or ethnicity on the principle of the secular and unitary nature of the French Republic. In fact, environmental justice issues are not likely to be perceived, analyzed and framed in racial and ethnic terms in Europe, but in terms of social categories (Laurent, 2011).

Table 2.3: Definition of variables

Variable	Definition	Sources
M	Non-incidental mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the triennial period used. Unit: per 100,000 people	National federation of regional health observatories (ORS)
Mr	Non-incidental male mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the triennial period used. Unit: per 100,000 people	National federation of regional health observatories (ORS)
Ms	Non-incidental female mortality rate, age standardized rates 2000-2004 calculated using data on registered deaths from INSERM, CEPIDc and INSEE. The year corresponds to mid-year of the triennial period used. Unit: per 100,000 people	National federation of regional health observatories (ORS)
NO_2, PM_{10},O_3	Annual mean of NO_2, PM_{10}, O_3 concentration (μg/m^3) respectively 2000-2004	French Environment and Energy Management Agency (ADEME)
Pr	High precipitation totals 2000-2004	Météo France
Sun	Annual accumulation of sunlight in hours 2000-2004	Météo France
Frost	Annual number of frost days in days 2000-2004	Météo France
Wind	Annual instantaneous maximum wind in km/hour 2000-2004	Météo France
Sm	Number of cigarettes sold per 1,000 residents 2000-2004	French Monitoring Centre for Drugs and Drug Addictions (OFTD)
Alcohol	Annual number of deaths related to alcohol. It includes liver cirrhosis, alcoholic psychosis and alcoholism, cancer of the upper aero-digestive tract	National federation of regional health observatories (ORS)
Accident	Road accident rate, age standardized rates 2000-2004 calculated using data on registered road accident. Unit: per 100 000 people.	National federation of regional health observatories (ORS)
PPHB	Number of people per 1 hospital bed 2000-2004	National federation of regional health observatories (ORS)
Industry	Share of industry in the total value added of a department (in %).	French National Institute for Statistics (INSEE), census 2005
Education	Population from 15 years (without students) with minimum BAC+2 divided by population within department in 2006	French National Institute for Statistics (INSEE), census 2007
Un	The unemployment rate is the percentage of unemployed people in the labor force (occupied labor force + the unemployed) 2000-2004.	French Ministry of Labor (DARES)
Income	Income is defined as the net taxable income divided by the number of taxed households within a department	French National Institute for Statistics (INSEE)
The intensity of poverty	(or poverty gap) is an indicator used to assess the extent to which the standard of living of the poor population is under the poverty line. It is calculated formally as follows:(poverty threshold - median standard of living of the poor population) / poverty threshold.	French National Institute for Statistics (INSEE), census 2005

The following panel describes the control variables. Data on weather come from Meteo France via the French Institute of the Environment (IFEN). The smoking rate fell by 35% between 2000 and 2004, probably due to the "Loi Évin" of 1991 prohibiting advertsing of cigarettes, and tax increases (INSEE). From the ORS I also collected data about the number of people per hospital bed, to measure the health care system and the availability of medical care resources in a particular department, from 2000 to 2004. I add the share of industry to control for industrialization,[4] as a time invariant variable for each department, and based here on the 2005 census by the French Institute of Statistics (INSEE).

Finally, the last row of descriptive statistics corresponds to an air pollution index. To assess the sensitivity of the weighted measure of pollution I have described above, I use the Atmo index, the air pollution index calculated by the AASQA. The Atmo outlook varies daily according to air quality using a scale of 1-10 (1 = very good air quality, 10 = very bad air quality). This index takes into consideration the concentration of four sub-indexes characterizing Nitrogen Dioxide (NO_2), Sulphur Dioxide (SO_2), Particles in suspension (PS) and Ozone (O_3). It considers pollution measured only by urban and industrial monitoring stations for main conurbations, for a period from 2000 to 2003. After 2003, the construction of the index was changed, so that I cannot consider it for 2004. I retained 41 conurbations and I associated each one with a department. I constructed a yearly variable summing up the number of days

4. French data about industrialization and GDP are not precise enough to take into account time fluctuations among departments from 2000 to 2004. The pourcentage of industry value added over the total value added for each department is available only every five years (INSEE).

above indices 8, 9 and 10, which corresponds to poor air quality according to the definition of the Atmo index (Table 2.4).

This variable is a proxy for peaks of pollution and is correlated positively with the weighted measure of air pollution concentration, which gives it more credence. This index variable is positively correlated with the previous measure of NO_2, NO, PM_{10} and O_3.

However, further in the estimation, I prefer to use real concentrations of pollution instead of indices. A few days correspond to peaks of pollution, and fixing a threshold below which pollution does not have any impact is highly arguable. Pollution does indeed fluctuate, a low level can be active and the level perceived as toxic is variable, even among the healthy population. Within a population, some people are more sensitive than others and will suffer from atmospheric pollution even at really low levels: levels below the actual threshold fixed by the public authorities. I aim to test this intuition.

Table 2.4: The Atmo index

Index scale	Sub-indexes	PM_{10} **scale** Average of mean daily concentrations in $\mu g/m^3$	NO_2 **scale**	O_3 **scale** Average of the hourly maxima in $\mu g/m^3$
Very good	1	0 - 9	0 - 39	0 - 29
Very good	2	10-19	40 - 79	30 - 54
Good	3	20 - 29	80 - 119	55 - 79
Good	4	30 - 39	120 - 159	80 - 104
Moderate	5	40 - 49	160 - 199	105 - 129
Poor	6	50 - 64	200 - 249	130 - 149
Poor	7	65 - 79	250 - 299	150 - 179
Bad	8	80 - 99	300 - 399	180 - 209
Bad	9	100 - 124	400 - 499	210 - 239
Very bad	10	125 and more	500 and more	240 and more

2.4. Model and Econometrics

2.4.1. Specification

The focus of this study is the relationship between average pollution, socioeconomic status, and mortality. The unit of analysis is the department, which is the main administrative unit in France, below the national and regional levels. Departments are thus administrative divisions. They form one of the three levels of local government, together with the 22 mainland and 5 overseas regions, in which they are grouped. There are 95 departments in France with an average population of 620,000 people, ranging from over 70,000 to over two million.[5]

In the analysis, I start by estimating a standard model with the non-incidental mortality rate as the dependent variable without considerations of environmental quality. After running preliminary regressions for various functional forms and following the results from an overall normality test based on skewness and on kurtosis for each, I estimate an equation of the following form to ensure that errors are distributed normally $\varepsilon \sim N(0, \sigma^2)$:[6]

$$X_{it}^k = \alpha_k + Socioeconomic_{it}\beta_k + Demographics_{it}\eta_k + Z_{it}\phi_k + \varepsilon_{it}^k \tag{2.1}$$

where i refers to local authority, t to the year, and k the kind of mortality rate. X_{it}^k

5. Due to missing data, some departments were withdrawn from the analysis in order to create a balanced panel, so that 41 departments were finally used.

6. There is no evidence that the log transformation is the best fit for mortality time trends (Bishai & Opuni, 2009). Moreover, given the size of the department, the effect of outliers may not be a problem here.

is a vector of non-incidental mortality rates (overall mortality rate, male and female mortality rate). Socioeconomic variables and particularly the unemployment rate and income are included as the main explanatory variables. Due to multicollinearity issues, I do not include both the average income and the education variable. The squared correlation between education and the average revenue is above 0.8.

The vector $Demographics_{it}$ includes several variables. First, it accounts for lifestyle, which refers to the regular activities and habits a person has that could have an effect on his or her health. I include the smoking rate and the alcohol variable as a proxy for lifestyle. Likewise, health effects for some of these pollutants are likely to be associated with the amount of time a person spends outdoors, even if I am not able to include it in the analysis. The vector also accounts for age structure across departments including a variable for the population more than 85 years old, and the population less than 4 years old. The magnitude of health effects is likely to be higher in areas with more young children or elderly, since they are particularly vulnerable. The number of people per hospital bed $PPHB_{it}$ in each department is included as a proxy to measure the health care system and the availability of medical care resources in a particular department. I also include the percentage of industry value-added over the total value-added $Industry_i$ for each department, as a time-invariant variable. I further take into consideration weather patterns at department level Z_{it}, as a control for average pollution levels. I consider the annual mean of precipitation to capture the effect of very wet years, the maximum wind speed, the number of frost days, and the annual accumulation of sunlight, as a time varying control. Some studies contend that mostly long-term (i.e., monthly and annual) fluctuations in temperature affect mortality (Martens, 1998). Besides, wind speed measurements are important for air quality monitoring. The higher the wind speed, the lower the pollutant concentration. Wind dilutes pollutants and rapidly disperses them throughout the immediate area. ε_{it} is the error term.

Independent variables explain most differences between departments and years, but there is probably some unmodeled heterogeneity. Thus, the next step in performing a multilevel analysis is to decide whether the explanatory variables considered in the analysis have fixed or random effects. The Hausman test considers the null hypothesis that the coefficients estimated by the efficient random effects estimator are the same as the ones estimated by the consistent fixed effects estimator. By running this test, the fixed effects model appears to be the most efficient one. In fact, I think of each department as having its own systematic baseline. I also calculate the robust variance estimator, in order to prevent the heteroskedasticity that I found by running Breush-Pagan test: this test checks if squared errors are explained by explanatory variables. The estimation will also take into account autocorrelation, because the Wooldridge test shows that disturbances exhibit autocorrelation, with the values in a given period depending on the values of the same series in previous periods. To address the possibility that omitted variables account for some of the heterogeneity among French departments, an error component model is estimated:

$$\varepsilon_{it} = c_i + \delta_t + u_{it} \tag{2.2}$$

c_i and δ_t are residual differences where c_i is a department effect which accounts for differences across departments that are time-invariant (e.g lifestyle differences that we cannot take into account), δ_t is a year effect which controls for factors that vary uniformly across departments over time, and u_{it} is the remaining error term.[7] It is also likely that a population's health affects unemployment via productivity, education and other factors This potential simultaneity can be a source of endogeneity, making standard estimators inconsistent. I need to test this hypothesis, so I consider the lag of the endogenous variable - unemployment - as an instrument. The F-test on the excluded instruments in the first stage regression confirms the validity of the instrument. To avoid the weak instruments pathology, we look at the F-test on the excluded instruments in the first stage regression and check whether the test statistic is greater than 10 ($F(1,192) = 28.91$). Then, the Hausman test rejects the endogeneity of the model ($P=0.810$). I also test income as an instrument. However, I find that income and the lag of income are not good instruments, respectively ($F(1,190)= 3.86$; $F(1,189)= 1.33$).

This paper is also concerned with spatial correlations which could bias the results or introduce inefficiency. If the observations are spatially clustered, the estimates obtained will be biased or inefficiency will be introduced. In fact, the mortality rate in one region could be related to that in another.[8] As a result, I calculate the Driscoll and Kraay non-parametric adjustment of standard errors model allowing

7. The Ramsey test confirms the robustness of the specification.

8. The Moran Index of spatial contiguity rejects the null hypothesis that there is no spatial clustering of the value in the raw mortality data: the first tail test yields $I=0.266$ with a 1 percent probability

for both space and time adjustments.

In the second model, the mortality rate is expressed as a function of environmental variables added to the previous variables. I estimate the folowing model:

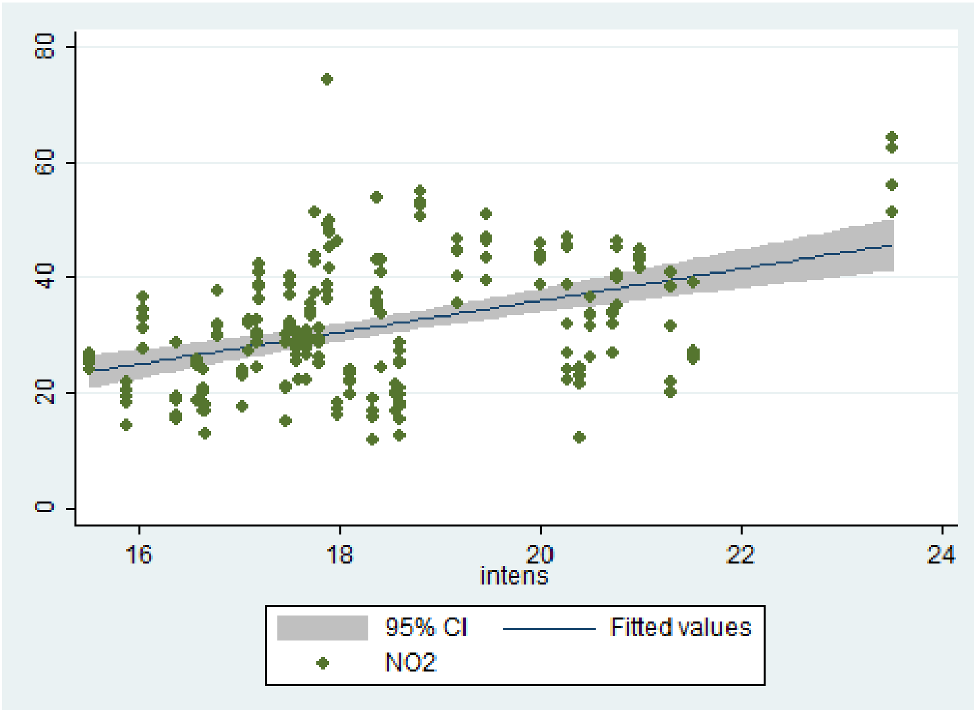
$$X_{it}^{k} = \lambda_k + P_{it}\theta_k + Socioeconomic_{it}\psi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^{k} \tag{2.3}$$

P_{it} is a vector of air pollutant concentrations for O_3, NO_2 and PM_{10}. In this model, the main coefficient of interest is θ, representing the mean parameter estimates for all 95 departments for explanatory variables P_{it}.. It also represents the effect of air quality on health outcomes. I again use the fixed effects estimator to control for heterogeneity between departments, with and without the Driscoll and Kraay standard errors model. Besides, the endogeneity problem has to be discussed in this context. I include fixed effects and some controls to address the problem of unobservable variables. However; there may be time-varying unobservable variables, not common to all regions and not captured by the dummies, which could bias the estimates. One may argue that European unemployment fluctuates around a very low level (Blanchard & Summers, 1986), making the previous Hausman test really weak. An association between the business cycle and mortality could, for instance, be driving the result (Chay & Greenstone, 2003). However, the French Statistical Institute does have access to the yearly business cycle data for each department. Another endogeneity bias could be that people may move in response to pollution levels. People who care more about health, and hence live a healthier life style, may move to less polluted areas, introducing an upper bias in the estimate of pollutants. However, migrations in France between departments are essentially in border areas and are mainly due to preferences for urbanization (INSEE).

If we think about environmental justice, the spatial distribution of pollution among departments may also be another reason why mortality rates in richer geographical areas are lower. Lower pollution in richer departments may influence positively their health, relative to departments with a high intensity of poverty. Figure 2.2 supports this idea, showing a potentially positive relationship between NO_2 and the intensity of poverty. Departments with high pollution levels seem to be those with a low socioeconomic levels. Secondly, poor people may also be more sensitive to pollution which threatens their health. To study this possibility, I estimate a model dividing my sample with those above and below the median of poverty gap as shown in Table 2.10. To do so, I create a dummy with respect to the index which is being considered: 50% of departments are below and 50% are above this median.

The intensity of poverty is an indicator used by INSEE to assess the extent to which the standard of living of the poor population lies below the poverty line. The higher the indicator, the greater the poverty gap is said to be, in that the standard of living of the poorest is a very long way below the poverty threshold. It may be asked if a change in pollution also benefits health more in areas with high socioeconomic levels than low levels areas. Are the health damages of an increase

Figure 2.2: Correlation between poverty gap and NO_2



in air pollution bigger in poorer areas compared to others? The next section shows that socioeconomic factors - in particular the unemployment rate - greatly interfer when studying the impact of NO_2 on mortality rates, at different levels of poverty.

2.4.2. Results

Impact of environment quality on health I start by examining a standard model for the mortality rate without consideration of environmental quality. I then add NO_2, O_3 and PM_{10} to the specification to see if considering pollutant variables improves the overall fit of the model. To capture the department effect and spatial autocorrelation, both fixed effects clustered at the regional level and the Driscoll and Kraay standard errors with fixed effects are estimated. Approximately 70 percent of the variation in the response variable may be attributed to explanatory variables.

Table 2.5: Standard model of non-incidental mortality rate

	OLS	FE	D-K S.E
Income	-0.0165***	-0.0641***	-0.0641**
	(0.00253)	(0.00596)	(0.0141)
Un	8.217*	17.01**	17.01***
	(4.435)	(7.432)	(2.593)
Pr	-16.83	3.755	3.755
	(9.914)	(5.898)	(5.744)
Sun	-0.0793***	0.0764***	0.0764**
	(0.0178)	(0.0174)	(0.0206)
Wind	0.786***	0.271**	0.271
	(0.267)	(0.120)	(0.199)
Frost	0.728**	-0.966***	-0.966*
	(0.290)	(0.209)	(0.425)
Sm	0.0619**	-0.105***	-0.105*
	(0.0231)	(0.0247)	(0.0412)
Industry	3.008**		
	(1.404)		
PPHB	0.0317	1.690	1.690
	(0.237)	(2.411)	(0.876)
Alcohol	0.0840	0.174*	0.174**
	(0.0508)	(0.0901)	(0.0613)
department FE		x	x
Observations	203	203	203
R-squared	0.598	0.753	

Note: This table presents the standard model of the non-incidental mortality rate with its main determinants. All regressions are estimated with standard errors clustered at the regional level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

I first estimate the standard model with the OLS, trying to test the most complete model, and I observe in the first column from Table 2.5 that most of the coefficients of the determinants of mortality are significant. However, I also analyse within estimator as I assume that the unobserved factors c_i between departments determine both mortality rates and explanatory variables. The fixed effects imposes time-independent effects for each entity that are possibly correlated with the regressors, which is why $Industry_i$, a time invariant variable, is not taken into account. The last column shows the Driscoll and Kraay standard errors model which takes spatial autocorrelation into account, with fixed effects. I observe that income impacts the mortality rate negatively in every regression, at a 1% level of significance. I am in line with the literature saying that income is a significant determinant of health. Ettner (1996) finds that increases in income significantly improve mental and physical health. Inadequate education and living conditions ranging from low income to the unhealthy characteristics of neighborhoods and communities can harm health through complex pathways. Health disparities by income are partly explained by disparities in medical care. French departments with a high level of income are more likely to have a low mortality rate than departments with a lower income. As a consequence, a shortsighted political focus on reducing spending in education, child care, jobs, community and economic revitalization, housing, and transportation could actually increase medical costs by magnifying disease burdens and widening health disparities.

I then study the relationship between NO_2, O_3, PM_{10} and mortality rates in both a single pollutant model (Table 2.6) and in a multi-pollutant one (Table 2.7). The multi-pollutant model allows coefficients to be examined at the same time, so as to not overestimate the impact of one pollutant. In addition, the F-test on the joint significance of the three pollutants rejects the hypothesis that they are jointly insignificant (P = 0.8612).

As shown in the single pollutant model, coefficients for PM_{10} are positive and significantly different from zero with the Driscoll and Kraay estimation. When I con-

Table 2.6: A simple pollutant model of mortality

VARIABLES	(1) FE	(2) D-K S.E	(3) FE	(4) D-K S.E	(5) FE	(6) D-K S.E
NO_2	0.438*	0.438*				
	(0.237)	(0.182)				
O			0.220	0.220		
			(0.377)	(0.195)		
PM_{10}					0.193	0.193*
					(0.263)	(0.0881)
Income	-0.0645***	-0.0645**	-0.0646***	-0.0646***	-0.0649***	-0.0649***
	(0.00568)	(0.0142)	(0.00543)	(0.0134)	(0.00560)	(0.0134)
department FE	x	x	x	x	x	x
Weather controls	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x
Observations	203	203	203	203	203	203
Adjusted R-squared	0.756		0.754		0.754	

Note: This table presents the impact of NO_2, O_3 and PM_{10} on the non-incidental mortality rate. All regressions are estimated using fixed effects with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather patterns with wind, sun, precipitation and frost; a vector of socioeconomic variables including unemployment rate and income; and a vector of demographics including the level of industrialization, people per hospital bed and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

sider a multiple pollutant model, PM_{10} is significant and negative with the Driscoll and Kraay estimation. This may be explained by the high interaction between O_3 and NOx which may bias the result obtained for PM_{10}. Secondary particles are also formed from other pollutants such as NO_2 . Variation in the dataset may also not be sufficient to obtain significant results for PM_{10} even though the single pollutant model shows a significant and positive impact of PM_{10} on mortality rates. Mortality rates do not vary too much within the region over time. Moreover, we only have access to the variation within departments over a few years, from 2000 to 2004. This may show very little variation, perhaps some of it is due to measurement errors which would bias coefficients towards zero. This result is in line with Chay *et al.* (2003) who examine the effect of particulate matter on adult mortality, in the US during the 1970s. They find no impact of this source of pollution on adult mortality. Yet, this result is contradicted by the French study by the Sanitary Health Institute which found a positive effect of PM_{10} on mortality, in a panel of nine different French cities (Pascal *et al.*, 2009). However, this study does not specify the type of estimator used. Furthermore, Pascal *et al.* (2009) do not take into account the influence of lifestyle or socioeconomic factors on health, as their model strictly includes weather data, whereas the robustness of the model is not verified if I take socioeconomic factors out. Finally, the average concentration from the measure is probably lower than the level used by the Sanitary Health Institute which considers 9 urban cities.

Ozone is positively correlated with the mortality rate but it is not significant in both a simple and multiple pollutant model when using fixed effects. The relationship between Ozone and temperature remains a complex phenomenon and may be the cause of the lack of significance. It seems highy complex to untangle one effect from the other. The Sanitary Health Institute in France emphasizes the complexity of studying the interaction between Ozone and sanitary variables because temperature, humidity, winds, and the presence of other chemicals in the atmosphere influence Ozone formation, and the presence of Ozone, in turn, affects those atmospheric constituents. French data show this positive correlation between temperature and Ozone.

In contrast, NO_2 appears to have a significant and positive effect in both single and multi-pollutant models, when I consider the fixed effects regression model with Driscoll and Kraay standard errors. These results suggest NO_2 has a positive a significant impact on the mortality rate in France at a department level. The fixed effects estimate suggests that for every 2 $\mu g/m^3$ increase in NO_2, there is almost one more death a year per 100,000 persons.[9] Concentrations of NO_2 vary from 12 to 74 $\mu g/m^3$, suggesting a difference of nearly 20 deaths a year per 100,000, depending on the department. At high levels of pollution, this result confirms the existence of a short-term relationship between current air pollution levels and mortality in France.[10]

I also observe from the OLS estimator that the effect of NO_2 on the mortality rate tends to lead to erroneous conclusions, if the fixed-effects problems are neglected. As a result, I give more credence to fixed effects estimators and specially the fixed effects regression model with Driscoll and Kraay standard errors for the rest of the study. To deal with the strong correlation existing between the three pollutants, I replace pollution variables by the atmo index in the second block of the multiple pollutant model of mortality. However, the atmo index variable is not significative in any of the three models estimated. Atmo index may be not precise enough to highlight a pollution effect. As pollutant models do not show any significance or robustsness for PM_{10} and Ozone for the reasons I have explained previously, I focus on the unique pollutant, NO_2, as it has a relevant significance in both models. Besides, focusing on

9. The death are registered in the municipality of death
10. Short-term effects studies refers to daily variation whereas long-term studies use cohort studies over several years

one pollutant allows coping with the high correlation existing between pollutants. In other words, NO_2 may be a proxy for several pollutant.

Table 2.7: A multiple pollutant model of mortality

VARIABLES	(1) OLS	(2) FE	(3) D-K S.E	(4) OLS	(5) FE	(6) D-K S.E
NO_2	0.261	0.611*	0.611**			
	(0.596)	(0.359)	(0.180)			
PM_{10}	0.0990	-0.285	-0.285*			
	(0.505)	(0.424)	(0.124)			
O	-0.169	0.272	0.272			
	(0.535)	(0.337)	(0.167)			
atmo				-0.665	-0.542	-0.542
				(1.104)	(0.460)	(0.330)
Income	-0.0165***	-0.0645***	-0.0645**	-0.0215***	-0.0891***	-0.0891***
	(0.00218)	(0.00560)	(0.0143)	(0.00368)	(0.0103)	(0.0103)
Constant	901.8***	1,518***	0	995.1***	1,694***	0
	(121.3)	(254.1)	(0)	(152.0)	(319.5)	(0)
department FE		x	x		x	x
Weather controls	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x
Observations	203	203	203	119	119	119
Adjusted R-squared	0.745	0.581		0.619	0.811	

Note: This table presents the impact of a multiple pollutant model on the non-incidental mortality rate . All regressions are estimated using fixed effects with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather patterns with wind, sun, precipitation and frost; a vector of socioeconomic variables including the unemployment rate and income; and a vector of demographics including the level of industrialization, people per hospital bed and the smoking rate. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Gender Analysis I now consider separately female and male mortality rates as dependent variables related to NO_2. Figure 2.1 presents the relatively stable distribution of the non-incidental mortality rates by gender, as well as the overall mortality rate. Ms. corresponds to female mortality rate whereas Mr. corresponds to male mortality rate. Results are detailed in Table 2.8. Income is again negatively and highy correlated with the female and male mortality rates for both models. I observe a significant effect for females as I do not find any for males. I am in line with the previous literature (Clougherty, 2010) (Marr, 2010): women are more sensitive and vulnerable to pollution. Smoking rates and unemployment have a significant and positive effect on male mortality rates when considering fixed effects models. In contrast, the female mortality rate model shows significativity for the pollution variable but not for the smoking rate. Note that the regression for the male mortality rate has a higher R-squared than for the female mortality rate.

The female fixed effects estimate suggests that for every 2 μg/m^3 increase in NO_2, another death of a women is registered, suggesting a difference of 30 deaths a year for women across departments. Lifestyle is represented here with the smoking rate and alcohol seems to be more significant than air pollution concentrations for male mortality rates. Moreover, the relative impact of alcohol and having a job on health seems to be greater for men. This may also be a case of the ecological fallacy described previously. Even if at the individual level there is negative correlation between unemployment and pollution, there may be a positive correlation at the aggregate level. This result raises questions about the implications for individuals' levels of exposure relative to demographics, types of activity or personal health situations.

Table 2.8: A gender model of mortality

	female Mortality rate		male Mortality rate	
VARIABLES	FE	D-K S.E	FE	D-K S.E
NO2	0.566***	0.566***	0.00910	0.00910
	(0.181)	(0.0760)	(0.242)	(0.145)
Income	-0.00540*	-0.00540	-0.0254***	-0.0254***
	(0.00294)	(0.00284)	(0.00491)	(0.00322)
Sm	-0.0114	-0.0114	0.0135	0.0135
	(0.0120)	(0.00873)	(0.0140)	(0.00635)
PPHB	0.471	0.471	1.598	1.598**
	(1.396)	(0.514)	(2.765)	(0.498)
Un	0.881	0.881	4.654	4.654***
	(3.103)	(0.669)	(2.905)	(0.494)
department FE	x	x	x	x
Weather controls	x	x	x	x
Demographic controls	x	x	x	x
Socioeconomic controls	x	x	x	x
Observations	203	203	203	203
Adjusted R-squared	0.402		0.764	

Note: This table presents the impact of a multiple pollutant model on the non-incidental mortality rate . All regressions are estimated using fixed effects with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather patterns with wind, sun, precipitation and frost. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Interaction between socioeconomic status and environment quality Furthermore, I might suspect that the effects of exposure to air pollution on health varies with socioeconomic status, in France. People with low incomes may be disproportionately exposed to environmental contamination that threatens their health.

First, the previous models in this paper show that income is negatively and very significatively related to mortality rates. Richer departments have a lower mortality rate. However, this study does not answer precisely the reasons behind this result. Better access to health care may be a reason why mortality rates in richer geographical areas are lower. There is a positive correlation between the level of income and people per hospital bed in the dataset.[11] Once again it may also be the case of ecological inference. This negative correlation at the aggregate level may be positive at the individual level.

If we think about environmental justice, the spatial distribution of pollution among department may also be another reason why mortality rates in richer geographical areas are lower. Lower pollution in richer department may influence positively their health relatively to departments with a high intensity of poverty. Table 2.9, which present average values of mortality rates above and below the median of income supports this idea.

Secondly, poor people may also be more sensitive to pollution which threatens

11. The simple coefficient of correlation indicates a positive correlation of 0.2 between income and people per hospital bed within the French departments.

Table 2.9: The average values of mortality rates for different thresholds of pollution and socioeconomic status

Variable	Mean	Std. Dev.	Min.	Max.	N
	Low income(below the median)				
non-incidental mortality rate	781.8545	67.2788	670	1000	110
	High Income (above the median)				
non-incidental mortality rate	745.7455	68.35437	584	915	110
	Low pollution(below the median)				
non-incidental mortality rate	773.9455	62.39896	670	953	112
	High pollution (above the median)				
non-incidental mortality rate	753.6545	75.86364	584	1000	108

their health. To study this possibility, I divide the panel into departments above and those below the median of the intensity of poverty as shown in Table 2.10. The intensity of poverty is an indicator used by INSEE to assess the extent to which the standard of living of the poor population is under the poverty line. The higher the indicator, the greater the poverty gap is said to be, in that the standard of living of the poorest is a very long way below the poverty threshold.

The first two columns of Table 2.10 present the results for the sample above the median of poverty gap compared to the last two columns of the sample below. I observe that the coefficient for income remains highly significant and negative in both samples suggesting again the existence of health disparities with respect to income. The size of the impact is similar for both samples. The level of income does not have a higher impact on the mortality rate when considering departments with a high poverty gap with respect to their counterparts. We cannot conclude there is a poverty trap effect from this result. I observe a similar result for the impact of NO_2 concentration on mortality rate. Besides, I note that the sample with a low intensity of poverty seems slightly more affected by pollution than the sample above the median of poverty gap. This result is opposed to the intuition.[12] Coefficients difference for NO_2 may come from the unemployment variable.

Unemployment seems to moderate the effect of NO_2 on the mortality rate. The coefficient for unemployment is indeed positive and significant in the sample above the median of the intensity of poverty in the fixed effects model, whereas it is not significant within departments with a lower poverty gap. The size of the coefficient for unemployment in the sample of departments with above the median level of poverty is higher than the size of the coefficient for unemployment in the sample above. This result suggests that the impact of unemployment on mortality rates relative to pollution is more significant for departments with high poverty gaps than for departments with low poverty gaps. Unemployment tends to make the effect of NO_2 on mortality rates disappear. The poorer a department is, the more the effect of unemployment relative to pollution is manifest. The variable of interaction between NO_2 and unemployment presented in Table 2.11 confirms this intuition.

12. However, confidence intervals overlap with each other suggesting both coefficients are not significantly different.

Table 2.10: A model with respect to poverty gap

	(1)	(2)	(3)	(4)
	high poverty gap		low poverty gap	
VARIABLES	FE	D-K S.E	FE	D-K S.E
NO2	0.165	0.165***	0.387**	0.387*
	(0.120)	(0.0319)	(0.176)	(0.144)
Income	-0.0134***	-0.0134***	-0.0117**	-0.0117*
	(0.00348)	(0.00158)	(0.00454)	(0.00466)
Un	4.725*	4.725	1.671	1.671
	(2.578)	(2.279)	(5.463)	(1.093)
Department FE	x	x	x	x
Weather controls	x	x	x	x
Demographic controls	x	x	x	x
Observations	94	94	109	109
Adjusted R-squared	0.629		0.444	

Note: This table presents the impact of a multiple pollutant model on the non-incidental mortality rate with respect to the level of poverty . All regressions are estimated using fixed effects with standard errors clustered at the regional level or with Driscoll and Kraay standard errors. I include in all estimations a vector of weather patternw with wind, sun, precipitation and frost; and a vector of demographics including people per hospital bed and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Robustness check I perform several robustness tests to make sure the estimation is not biased by any unobserved factors. In all the robustness estimations summarized in table 2.11, I take spatial autocorrelation and department effect into account, using the Driscoll and Kraay standard errors model with fixed effects.

First, to be more precise, I want to analyze whether the effect of the socioeconomic variables is moderated or modified by the introduction of the environmental variable. To do so, I include an interaction variable to look at how unemployment and NO_2 interact. I add an interactive term $P_{it}Socioeconomic_{it}$ between unemployment and environmental quality to provide a better description of the relationship between mortality rate and the independent variables such that:

$$X_{it}^k = \lambda_k + P_{it}\theta_k + Un_{it}\psi_k + P_{it}Un_{it}\varpi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^k \quad (2.4)$$

$$= \lambda_k + (\theta_k + Un_{it}\varpi_k)P_{it} + Un_{it}\psi_k + Demographics_{it}\sigma_k + Z_{it}\phi_k + \varepsilon_{it}^k$$

where $(\theta_k + Un_{it}\varpi_k)$ represents the effect of environmental quality on the mortality rate at a specific level of socioeconomic variables and ϖ_k indicates how much the slope of P_{it} changes as the unemployment variable goes up or down by one unit. To ease the interpretation, I consider a dummy variable for NO_2. NO_2 will take the value of "1" when departments are above the median of NO_2 concentration and "0" otherwise.[13] The coefficients for unemployment, NO_2 and the interactive variable are significant with the endogenous variable "non-incidental mortality rate". The significativity of the interaction coefficient suggests that the effect of NO_2 has been modified by the unemployment variable. In other words, the effect of NO_2 $(\theta_k + Socioeconomic_{it}\varpi_k)$ at some value of unemployment $Socioeconomic_{it}$ has a significant effect on the mortality rate. And the negative sign indicates the introduction of unemployment moderates the effect of pollution on mortality which confirms the intuition described above.

Besides, interacting income and pollution does not show any significativity. Sec-

13. Critics assert that an increased level of collinearity in models including a multiplicative term distorts the beta coefficients. However, a fixed effects model, or a mean purged regression model, automatically reduces multicollinearity.

ond, to deal with potential non linearity in the model, income squared is added to the estimation. The really small coefficient is not surprising knowing the average of 234.000.000. It may explain why the impact of income on mortality rate is not changing in the sample above and below the intensity of poverty. It also may be why there is no significativity when interacting pollution and income. Third, I include education to the estimation to make sure this variable does not biased results. There is also, most likely, a direct positive effect of education on health (Groot & Maassen van den Brink, 2007). While the exact mechanism underlying this link is unclear, the differential use of health knowledge and technology is almost certainly an important part of the explanation. We cannot conclude on the sign of the coefficient for education because there is a high correlation between income and education. However, the impact of NO_2 on mortality rate does not change neither in magnitude and in significativity when adding education to the model. Fourth, the introduction of population within the regression shows that population among department is not an issue. Fifth, I include a variable of pollution squared in order to compare results when facing higher atmospheric pollution. The positive and highly significant coefficient for NO_2 squared indicates that the impact of NO_2 on the overall mortality rate is amplified with higher pollution level. Finally, all over the paper, I have considered NO_2 as a measure of exposure to air pollution which was calculating using the annual measure of concentration weighted by the inverse of the distance between the census block where there is the most significant population and the monitoring station. As a last robustness check, I estimate a model considering the concentration of NO2 non weighted. Model 8 shows it does not change the result if I consider a measure without any weight. The impact of NO_2 is still positive and significant.[14]

14. However I prefer the weighted measure of concentration as I have described in the previous section because it tends to lead to a more accurate measure of exposure to air pollution.

Table 2.11: Robustness check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NO2	0.263**	0.966	1.487**	0.277***	0.340***			
	(0.0749)	(0.933)	(0.397)	(0.0600)	(0.0729)			
$Income^2$	-1.01e-06**							
	(2.66e-07)							
NO2*Income		-4.17e-05						
		(6.53e-05)						
NO2*Un			-0.122**					
			(0.0315)					
Education				46.23***				
				(7.713)				
Pop					-0.000329**			
					(0.000113)			
$NO2^2$						0.00516**		
						(0.00116)		
Non weighted NO2							0.649**	
							(0.191)	
department FE	x	x	x	x	x	x	x	x
Weather controls	x	x	x	x	x	x	x	x
Demographic controls	x	x	x	x	x	x	x	x
Socioeconomic controls	x	x	x	x	x	x	x	x
Observations	99	203	203	203	203	203	203	203

Note: This table presents the impact of a multiple pollutant model on the non-incidental mortality rate with respect to the level of poverty. All regressions are estimated using fixed effects and with Driscoll and Kraay standard errors, a vector of weather controls including wind, sun, precipitation and frost, a vector of socioeconomic variable including the unemployment rate and income and a vector demographics including people per hospital bed and and the smoking rate. Robust standard errors in parentheses. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

These results do not show greater effects among the more deprived. Thus, I cannot conclude to the existence of environmental injustice among French departments. Health disparities exist among French departments but seem to be more related to socioeconomic factors than differences in sensitivity to pollution. In fact, high level of unemployment tends to moderate the effect of pollution on health. However, the impact of pollution on mortality rates remain an important issue as we found a significant effect of NO_2 especially at levels below the current standard.

2.5. Conclusion

The objective of this paper has been first to investigate whether a department's environmental quality and socioeconomic status relative to its neighbors has an impact on its mortality rate. The second purpose has been to analyze the link between inequalities and air quality across departments. I test these hypotheses by using a multivariate model and taking spatial autocorrelation and fixed effects into account.

The results are strongly supportive of the hypothesis that NO_2 has a positive impact on mortality, with the effect being larger when considering higher level, of pollution. Moreover I show that even relatively low concentrations of air pollutants, at levels below the regulated threshold, are related to a range of adverse health effects. We may wonder about the need to reinforce NO2 standard.
I also shed light on the existence of health disparities in France. As a consequence, the choice of economic policies can have severe implications for health and medical

spending. Finally, I point out that women's health is more impacted than men's health by NO_2. This finding is consistent with the results of international studies that have examined the relationship between economic inequality, environmental quality and health. It also confirms the importance of ambient air pollution and reinforces the need for politicians to take into account environmental justice in France.

The paper suggests that further research on environmental inequality in France focusing on smaller geographical levels and individual characteristics is essential. It would be consistent to examine the impact of atmospheric pollution focusing on individual-level data. It would also be interesting to have access to morbidity data, especially for occupation-linked diseases, in order to shed light on the implications for losses of productivity.

2.6. Appendix

I assign annual pollutant concentrations to France's 95 departments. To do so, I assign a weight to every monitoring station. This weight corresponds to the distance between the census block of the department with the highest population and the monitoring station. This distance is given as the great-circle distance between two points, that is the shortest distance over the Earth's surface, or as the 'as-the-crow-flies' distance. Let (L_i,N_i) be the latitude and longitude in degrees of monitoring station i and (Lj, Nj) of the census block with the highest population j. The distance between the monitoring station i and the census block j is given by:

$$d_{ij} = arccos(G_{ij})R$$

where R is the radius of the earth, measured around the equator ($R = 6378$) and

$$G_{ij} = sin(aL_i)sin(aL_j) + cos(aL_i)cos(aL_j)cos(aN_j - aN_i)$$

with $a = \pi/180$ From this distance I calculate a weighted mean of pollutant concentration. The weighting attributed to a monitoring station corresponds to the inverse of the distance between the the census block with the highest population and the station so that every element C_{ij} of the distance matrix C is given by:

$$C_{ij}=\frac{1/d_{ij}}{\sum_{i=1}^{n} 1/d_{ij}}$$

Matrix C is a stochastic matrix of size N×N where elements in the main diagonal are zero. It is normalized in order to have each row summing to 1. Such normalization allows us to consider the relative distance instead of the absolute one. Then, I calculate the average weighted mean of pollutant concentration $\bar{P}$ within the entire department:

$$\bar{P} = \sum_{i=1}^{n} C_{ij} P_{iu}$$

P_{iu} corresponds to the annual mean concentration measured by the monitoring station i for the pollutant u.

Chapter 3

Energy Production and Health Externalities: Evidence from Oil Strike Refineries in France.[1]

1. This chapter is a joint work with Matthew Neidell from Columbia University

This paper examines the effect of energy production on newborn health using a recent strike that affected oil refineries in France as a natural experiment. First, we show that the temporary reduction in refining lead to a significant reduction in sulfur dioxide (SO_2) concentrations. Second, this shock significantly increased birth weight and gestational age of newborns, particularly for those exposed to the strike during the third trimester of pregnancy. Back-of-the-envelope calculations suggest that a 1 unit decline in SO_2 leads to a 196 million euro increase in lifetime earnings per birth cohort. This externality from oil refineries should be an important part of policy discussions surrounding the production of energy.

3.1. Introduction

Meeting the continued increased demand for energy is a major issue faced by nearly all countries. While there is much interest in developing renewable sources of energy, oil remains the predominant source given its relative price. Its portability also makes it particularly attractive for mobile sources, suggesting a reprieve in energy demand is unlikely in light of the tremendous growth in automobile ownership and travel throughout the world. Despite the price advantage of oil, its production poses a health risk. The point source emissions include several pollutants linked with numerous health impacts, most notably sulfur dioxide (SO_2). In some countries, such as France, nearly 20 percent of ambient SO_2 emissions come from oil production (Soleille, 2004). Evidence links SO_2 with a wide range of respiratory effects, and as such is regulated under environmental policies throughout the world.

The optimal design of energy policy must consider this production externality when comparing its full costs to those from renewable energy production. In this paper, we estimate the health effects from oil production by exploiting the pension reform strikes in France in October, 2010 that lead to a major disruption in the production of oil. These strikes provide an ideal natural experiment for overcoming the typical biases that arise when estimating the health effects of pollution. Amid nationwide protests over pension reform that involved raising the retirement age, striking workers blocked fuel supplies to oil refineries, which resulted in a complete cessation of operations at several major refineries for nearly a month. As Figure 3.4 demonstrates, this lead to a sharp reduction in SO_2 in areas close to the refineries when the strike began, that quickly dissipated once the strike was resolved and production resumed, while areas far from the refineries experienced no change in SO_2 levels. We exploit this temporal event by estimating difference-in-differences models, using areas far from the refineries as a control group. We focus on the health of newborns as the outcome of interest, both because this is a particularly sensitive group with much policy interest and because birth outcomes are strong predictors of a wide range of future outcomes (Black *et al.*, 2007) (Currie *et al.*, 2009).[2]

While this is not the first pollution-health study to use the closing of an industrial process or other exogenous event as a natural experiment[3], there are several

2. As noted in Joyce & Goldman (1988a) and Chay & Greenstone (2003), focusing on infants also offers a methodological benefit because cumulative exposure can be readily assigned, circumventing issues around mobility and prior exposure.

3. While there are a wide range of studies on this topic using quasi-experimental techniques (see the review in Zivin & Neidell (2013)), the most closely related are Ransom & Iii (1995), Hanna & Oliva (2011),

important features of our design that make this an important contribution, mostly centered on parameter identification. First, a common concern in such analyses is that individuals sort into residential locations based, in part, on the amount of air pollution and the employment opportunities in the area, making pollution exposure an endogenous variable.[4] A permanent change in pollution levels can lead to a temporary disequilibrium in the housing market whereby there is no sorting at the time the shock occurs, but sorting is likely to resume as time from the shock passes. If the "post-shock" period includes a long enough time period, then sorting, and hence the endogeneity of pollution, remains a potential concern. In our case, the closure of the refineries was a temporary event - lasting approximately one month - making it unlikely that households relocated in search of new employment opportunities or because of preferences for air quality. Second, seemingly exogenous events, such as a strike, may lead to unobserved behavioral changes in the treatment group that affect health, potentially invalidating the research design. Two features make this unlikely in our setting. One, although the variation in pollution is due to the closure of refineries at specific locations, the strike that caused this was a nationwide one centered on pension reforms, with the oil refineries an "unlucky recipient" of the protests. Therefore, any common responses to the strike are accounted for by including a control group. For example, changes in time allocation or activity choice because of the strike affected not only refinery workers but nearly all workers

and Currie *et al.* (2013), who all focus on the closing of industrial processes.

4. The link between employment opportunities and pollution endogeneity arises because industry creates both jobs and pollution.

throughout the country.[5]

Third, studies that examine the effect of prenatal insults often seek to uncover the distinct effects from different stages of the pregnancy in order to encourage the optimal use of prenatal care. In particular, shocks that occur early in pregnancy, specifically for women who are not yet aware they are pregnant, may leave little opportunity to engage in health-promoting behaviors (Almond & Currie, 2011). In the case of pollution, relatively simple behaviors, such as altering the amount of time spent outside, can yield significant improvements in health (Neidell, 2009).[6] Reliably estimating the separate contribution from each trimester is complicated by the fact that pollution levels are often highly correlated across the three trimesters of pregnancy, resulting in severe multicollinearity. Because the strike led to a sharp decrease in pollution for roughly one month, upon which it returned to baseline levels almost immediately after, our research design allows us to overcome this multicollinearity concern to more precisely investigate the separate effects by trimester. Lastly, the handful of quasi-experimental economic studies examining the impact of emissions from energy sources typically focus on the consumption of energy (Currie & Walker, 2011),(Beatty & Shimshack, 2011), (Moretti & Neidell, 2011), (Schikowski *et al.*, 2008). While this consumption side represents an important externality, the

5. Note that this strategy does not account for avoidance behavior, i.e., changes in time allocation in direct response to the changes in pollution ((Neidell, 2009)). This does not introduce a bias per se but changes the interpretation of estimates, so that our estimates reflect the effect of the strikes net of avoidance behavior. SeeZivin & Neidell (2013) for more details.

6. For example, air quality alerts, which seek to warn the public of dangerous air quality levels, are particularly targeted at pregnant women.

production externality is empirically distinct, but has received limited attention.[7] More reliable estimates of the health impacts from energy production are an important component in the development of policies surrounding energy production (Parry & Small, 2005) and the siting of industrial plants. Using this natural experiment, we first demonstrate that although SO_2 is considerably higher in areas close to the refineries, it falls significantly during the strike compared to areas far from the refineries, with regression results supporting the pattern in Figure 3.4.

We find no evidence of changes in two other pollutants, particulate matter and nitrogen dioxide, around the time of the strike, a finding consistent with the change in SO_2 coming from the oil refineries. Turning to health outcomes, we find that birth weight and gestational age of newborns living in the same census tracts as the refineries increased by over 3 and 1.5 percent, respectively, during the strike. Nearly all of the improvement in weight gain can be attributed to the increase in gestation. Furthermore, these effects are primarily driven by exposure during the third trimester of pregnancy, a time when most fetal weight gain occurs. Overall, our estimates suggest that the effects from oil production that accrue to newborns alone are quite sizeable and should be an important part of policy discussions surrounding the production of energy.

7. Furthermore, the common pollutants from energy consumption are carbon monoxide and particulate matter.

3.2. Background: Refineries, Air pollution and Health

3.2.1. Pollution and the refinery closure

Refineries are responsible for 20 percent of SO_2 release in France (Soleille, 2004). Oil refineries convert crude oil to everyday product like gasoline, kerosene, liquefied petroleum. Crude oil contains relatively high quantity of sulfur, which leads to the creation of sulfur dioxide when crude oil is heated at the refinery to produce fuel. The refining process also releases a large number of chemicals such as benzene, chromium and sulfur acid into the atmosphere, which limits our ability to conduct a proper instrumental variable analysis. France has 11 refineries that produce 89 million tons of petrol every year. The main 4 refining companies operating in France are Total, Shell, Esso and Ineos, located in the regions of Haute Normandie, Provence Alpes Côtes dAzur, Rhône-Alpes, Nord-Pas-de-Calais, Pays-de-la-Loire, Ile de France and Alsace. Total refineries are allowed to emit up to 3,500 tons of sulfur dioxide per year which corresponds to 9.6 tons a day. Due to protests over pension reform, protesters successfully ceased production in October, 2010 by mass picketing and the creation of physical blockades around fuel depots. As a result, production was reduced to a minimum or completely ceased for nearly 18 days until the strike was resolved. Closing a refinery is a complex process that requires anywhere from 2 days to one week according to the size of the refinery, and a comparable time period to re-open. Thus, the reduction in SO_2 is likely strongest between mid October and the beginning of November. We focus on the 4 refineries that completely shut down as a result of the strike.[8]

3.2.2. *Pollution and health*

Sulfur dioxide (SO_2) is one of a group of highly reactive gasses known as oxides of sulfur (SOx). The largest sources of SO_2 emissions are from fossil fuel combustion at power plants and other industrial facilities (EPA 2011). SO_2 is a colorless gas with a very strong smell. In France, the threshold for SO_2, fixed by the European Act of 2002-13 related to air quality, is 132 parts per billion (ppb) per hour; violations occur when this standard is exceed more than 24 times a year. In comparison, the Clean Air Act in the United States set the one-hour SO_2 standard at 75 ppb, where a violation occurs if the 99th percentile of 1-hour daily maximum concentrations, averaged over 3 years, exceeds this value. This standard was recently strengthened in June 2010, suggesting the need for reliable estimates of the relationship between SO_2 and health. Given the rapid stages of development that a fetus goes through in a short period of time, negative shocks can results in both immediate and latent effects (Almond & Currie, 2011). Pollution is one potential shock because it can impair the health of the mother, indirectly compromising fetus health, or cross the placenta, directly affecting the health of the fetus. Slama et al. (2008) describe more extensively possible biological mechanisms by which air pollutants may affect birth outcomes: SO_2, in particular, can harm the fetus by impacting blood viscosity and endothelial function. These changes can affect placental blood flow, transplacental

8. These refineries are Donges, Feyzin, Gonfreville l'Orcher and Petite Couronne.

oxygen and nutrient transport, all of which may affect fetal health. Furthermore, while there is a growing consensus that prenatal exposure to pollution affects birth outcomes, there is little understanding about the most susceptible periods of prenatal exposure. While the fetus experiences important organ developments in the first trimester, suggesting a particularly vulnerable stage, the fetus also gains the most weight during the third trimester, suggesting another crucial stage. Evidence from the fetal origins hypothesis suggests that exposure to negative shocks during early pregnancy has no effects at birth but latent impacts later in life (Almond *et al.*, 2009), while exposure during late pregnancy is more likely to affect birth outcomes (Stein *et al.*, 2003), (Schulz, 2010). Consistent with this, Deschenes *et al.* (2009) find that the sensitivity of birth weight to temperature is concentrated almost entirely in the second and third trimesters of the pregnancy. Whether these same patterns hold for pollution is largely unknown. While not focused on SO_2 per se, several economic studies have found robust evidence that prenatal exposure to pollution affects infant health (Currie *et al.*, 2009), (Sanders & Stoecker, 2011), (Currie & Walker, 2011). While most of these studies focus on the effect from exposure during the entire pregnancy, an important contribution of our study is the ability to precisely estimate the effects from exposure during each trimester. Furthermore, previous studies typically focus on pollution stemming from vehicular or industrial emissions, such as particulate matter and carbon monoxide, and our focus on oil refining is more relevant for SO_2.

A widely held position is that spending resources to reduce the incidence of low birth weight (LBW) births represents one of the best investments in improving infant health and welfare. For example, low birth weight has been linked to future health problems and lower educational attainment (Currie *et al.*, 2009). However we know from Almond *et al.* (2002) that birth weight cannot reliably substitute for direct health outcomes of interest when evaluating and predicting the effects of health policies and interventions. It may provide a misleading proxy for measuring the effectiveness of health interventions infant health outcomes. In addition, ud Din *et al.* (2013) shows that babies with low birth weight (LBW) catch up during adolescence. In addition, birth weight is a proxy of a wide number of elements: It may also underlie genetic conditions, childhood endowments. The economic valuation may then represent an upper bound of the cost of pollution. Difference in birth weight may be due to difference in cultural environment.

3.3. Data and empirical strategy

3.3.1. Data sources

Health data are drawn from the French National Hospital Discharge Database (PMSI) from 2007 to 2011. The key variables for our analysis are the year and month of birth, the census tract of residence of the patient[9], and the birth weight and gestational age at birth. Panel A of table 3.1 shows the birth weight and the gestational age by month, year and census tract. We also consider low birth weight

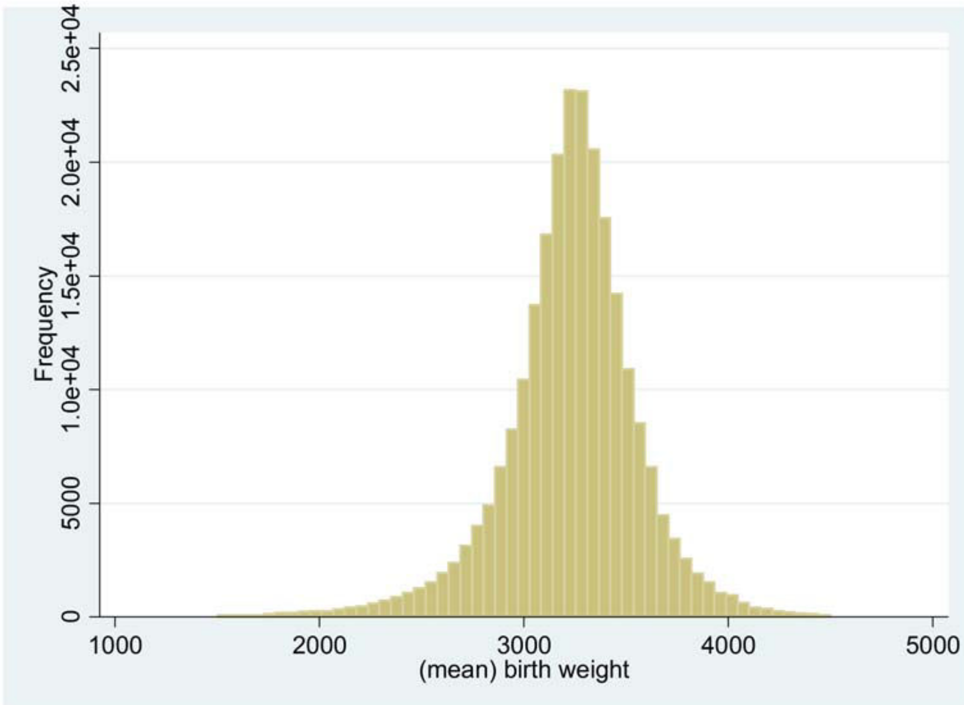
9. France have about 36600 census tracks

(<2500 grams) and short gestational age (<37 weeks) as two clinically relevant outcomes. We observe from table 3.1, panel A that the birth weight and gestational age are lower in census tracts with refineries (the treatment group) than in census tracts without refineries (the control group) for all periods of the study, hinting at potential effects from living near a refinery. Figure 3.1 shows the distribution of birth weight. Unlike the US, there is much less variation in birth weight in France, a finding consistent with universal access to health care. Birth weight after the refinery closure in the treated group has increased from 3207 to 3279 grams whereas birth weight in the other municipalities, has decreased from 3228 to 3320 grams in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is about 81.46 grams. Gestational age after the refinery closure in the treated group has increased from 38.70 to 39.08 weeks whereas gestational age in the other municipalities has decreased from 38.86 to 38.85 weeks in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is about 0.39 weeks.

Air quality is monitored throughout France by 38 approved air quality monitoring associations (AASQA). The French monitoring station system has approximately 700 measurement monitors equipped with automatic instruments. Figure 3.2 shows the location of monitoring stations, departmental boundaries (one of the three levels of government below the national level, between the region and the commune), and major cities throughout France.

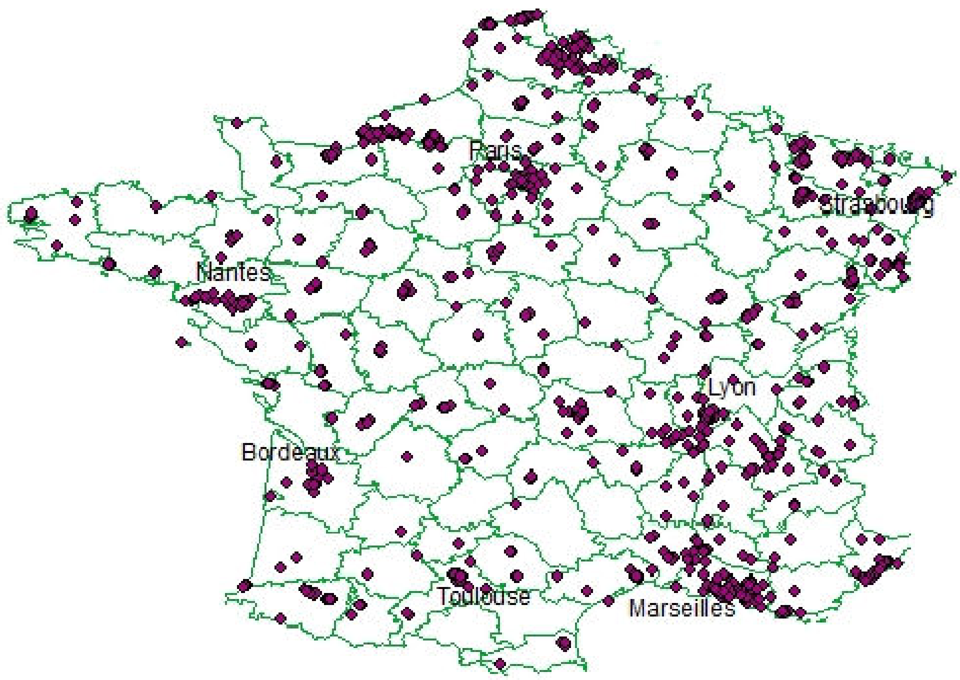
Not surprisingly, monitors are more highly clustered in major cities. The moni-

Figure 3.1: The distribution of birth weight



tors also show broad coverage of the country, with nearly every department having at least one monitor. We obtain daily measure of ambient air pollution concentrations in microgram per cubic meter (μg/m^3) for all air quality monitors in France for 2007-2010 from the Ministry for Ecology, sustainable development and spatial planning (ADEME) database. We also know the exact geographic location of each monitor. Since our main focus is on SO_2, we only include monitors that continuously measured SO_2 during this time period. This leaves us with 187 monitors that span 57 departments and 2864 census tracts. Monthly pollution concentration data are presented in Panel B of Table 3.1. The most notable aspect of this panel is that SO_2 levels are nearly 4 times higher in areas near the refinery, while the levels are

Figure 3.2: Air quality monitoring stations and department boundaries in France



virtually identical for PM_{10}, NO_2 and slightly higher for benzene.[10]

Sulfur dioxide concentration after the refinery closure in the treated group has decreased from 21.52 to 6.42 $\mu g/m^3$ whereas SO_2 concentration in the other municipalities has decreased from 1.03 to 1.65 $\mu g/m^3$ in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is 15.72 $\mu g/m^3$.

We also present the fraction of days in which the values recorded at the monitors exceeded health standards for SO_2 and PM_{10}.[11] While the number of exceedances is

10. Note that we dropped one inexplicably high measure of benzene (18.44) in order to make the scale of Figure 3.5 (below) easier to interpret. This measure occurred in a treated census tract on September 25, 2011, so including it would further reinforce the idea that the refineries may affect benzene levels as well.

11. There is no 24 hour air quality standard for benzene and NO_2. Although there is an hourly standard for NO_2, we were only able to obtain daily data.

quite low for SO_2 (occurring less than 1% of the time), census tracts with refineries are nearly 10 times more likely to have a violating monitor, consistent with higher SO_2 levels. The rate of exceedances for particulates is much higher on average, occurring nearly 5% of the time, though the rate of violations is quite similar across areas. Since weather has direct effects on health and also affects pollution formation, we also include meteorological data in our analysis. Our weather data come from Meteo France, the French national meteorological service. There are 100 monitors, one in each department. We also have daily measures at each monitor, along with data on the geographic location. We use average and maximum temperature, precipitation, maximum speed wind, prevailing wind direction, and maximum and minimum relative humidity. Summary statistics for daily and monthly measures of weather are presented in Panel C of Table 3.1. Although we include census tract fixed effects in our regression, which controls for all time invariant characteristics, we also include one measure of economic well-being to capture time varying factors: the unemployment rate. We use the quarterly rate of unemployment from the National Institute of Statistics and Economic Studies, which is available at the census tract level. Panel D of Table 3.1 also presents summary statistics for this variable.

Table 3.1: Summary statistics

	All	Treatment group (all time periods)	Control group (all time periods)	october 2010	Before and after oct2010
Outcomes					
birth weight (grams)	3228 [353]	3220 [272]	3228 [354]	3226.49 [355.16]	3229.84 [355.30]
birth weight < 2500 grams	.03 [.17]	.02 [.13]	.03 [.17]	.0322563 [.17]	.03 [.17]
gestational age (weeks)	38.86 [1.50]	38.78 [1.31]	38.86 [1.50]	[38.83 1.40]	38.85 [1.54]
gestational age < 37 weeks	.08 [.27]	.08 [.27]	.08 [.27]	.08 [.27]	.08 [.27]
Pollution					
SO_2 - monthly average ($\mu g/m^3$)	3.82 [4.53]	12.87 [10.86]	3.63 [4.10]	3.13 [2.75]	4.33 [5.07]
SO_2 - % days exceeding AQS yearly	.16 [3.89]	0.04 [1.13]	.12 [.35]	0 [0]	.16 [3.89]
NO_2 - monthly average ($\mu g/m^3$)	24.22 [14.81]	23.35 [12.27]	24.23 [14.83]	26.85046 [14.26]	25.2494 [15.67]
PM_{10} - monthly average ($\mu g/m^3$)	22.55 [7.45]	22.52 [6.76]	22.56 [7.45]	22.53 [4.41]	22.50 [7.32]
PM_{10} - % days exceeding AQS yearly	15.83 [14.64514]	1.46 [18.40496]	15.71 [14.6103]	1.60 [7.461515]	15.70 [14.68]
Benzene - monthly average ($\mu g/m^3$)	1.86 [.80]	2.78 [.74]	1.54 [.25]	1.59 [.83]	2.26 [4.46]
Covariates					
mean temperature (°C)	11.85 [5.91]	11.74 [5.65]	11.86 [5.92]	10.46 [6.05]	12.03 [5.87]
max. temperature(°C)	16.57 [6.76]	16.34 [6.55]	16.57 [6.76]	15.03 [7.10]	16.77 [6.67]
precipitation (mm)	2.19 [1.44]	2.28 [1.30]	2.19 [1.44]	1.90 [1.48]	2.22 [1.41]
wind speed (m/sec)	6.94 [1.45]	6.96 [.97]	6.95 [1.45]	6.84 [1.20]	6.96 [1.47]
wind direction (wind rose)	208.4 [40.4]	205.8 [38.1]	208.4 [40.4]	200.8313 44.73]	209.64 [39.55]
min. humidity (%)	55.07 [12.03]	56.32[12.50]	55.07 [12.03]	55.22 [14.77]	55.05 [11.55]
max. humidity (%)	92.7 [4.13]	93.19 [4.02]	92.6 [4.13]	92.32 [4.30]	92.72 [4.08]
unemployment rate (%)	8.71 [2.22]	8.50 [1.10]	8.71 [2.22]	9.34 [2.11]	8.69 [2.22]

Note: Reported values are means with standard deviations in brackets. The number of observations is 151,624. The control group has 151597 observations and the treated group 27 observations. Air quality standard (AQS) for SO_2 is 0.04 ppm (105 $\mu g/m^3$) for every 24 hour period and for PM_{10} is 50 $\mu g/m^3$ for every 24 hour period.

3.3.2. Merging data

Using the exact location of pollution and meteorology monitors and the census tract of residence for the birth outcomes, we assign pollution to census tracts in a two-step procedure. When a census tract has a pollution monitor in it, we assign that pollution concentration to the census tract. When it does not, we assign pollution using an inverse distance weighted average (IDWA) of pollution, similar to Currie & Neidell (2004). To do this, we compute the centroid of each census tract, and then compute the distance from the centroid to each monitor within the department. We then take the weighted average of pollution measurements from all monitors within a certain distance from the census tract centroid, using the inverse of the distance as weights. We vary the cutoff distance to assess the sensitivity of our results to our assignment of pollution. Although we have a daily measure of pollution and meteorology, health outcomes are only observed at a monthly level. We begin by aggregating pollution and meteorology at a monthly level. Since we only know the month of discharge for newborns, and their average length of stay in the hospital is 5.5 days, we must approximate their date of birth, and thus exposure to the strike. We assume all births occurred on the 1st day of the month, and assign pollution and meteorology from the previous 9 months (we also assess the sensitivity of results by assuming the 15th of the month). For example, an infant discharged in November is born anywhere from October 25th to November 25th, and we assume the birth date is November 1. We then assign exposure to this infant as the mean for the months from February through October, breaking it into 3 month intervals for examining trimester effects.

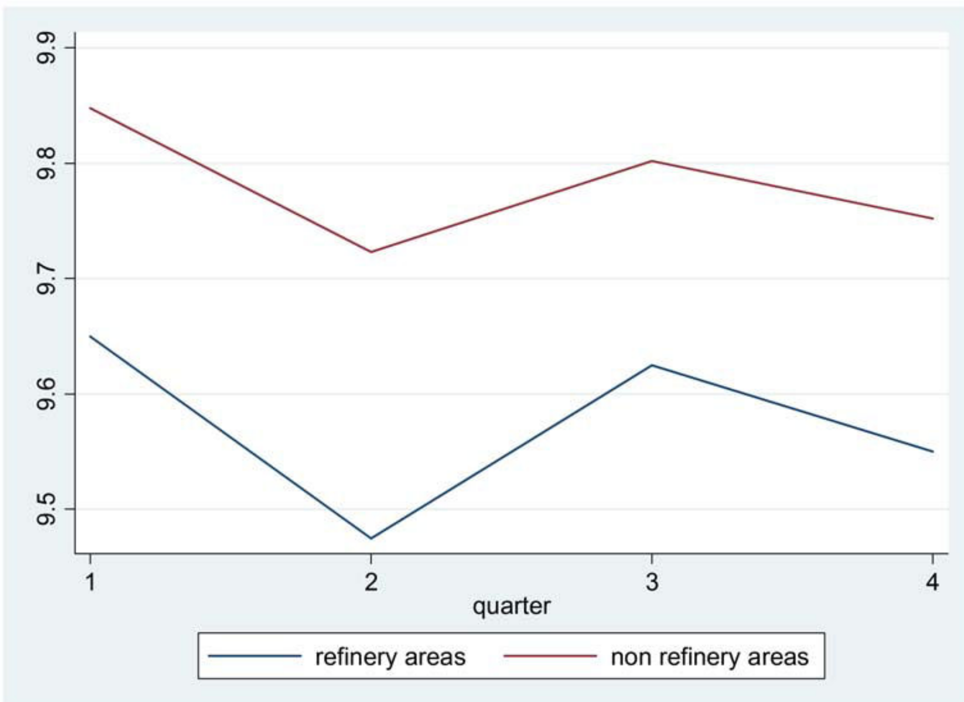
3.3.3. Empirical Methodology

Our goal is to assess the impact of oil production on both pollution levels and health outcomes at birth. We estimate difference in difference models to exploit the unexpected shutdown in production as a result of the strike in October 2010, using areas close to the refineries as the treatment group and areas far from the refineries as the control group. We implement this by estimating the following equation:

$$Y_{cm} = \beta * strike_m * close_c + \delta * X_{cm} + \sigma_m + \alpha_c + \epsilon_{cm} \tag{3.1}$$

where Y is either ambient pollution concentrations or birth outcomes in census tract c at month m. 'strike' is an indicator variable for the October 2010 period when the strike occurred, and 'close' is an indicator variable for whether the refinery is in the same census tract as the air pollution monitor or patient's residence. β is the difference-in-difference parameter. Xcm is a vector of census tract controls that include weather controls and the quarterly unemployment rate. We control for seasonal and temporal patterns by including month dummies and year dummies in σ_m. We include census tract fixed effects (α_c) to control for time-invariant characteristics of the census tract. ϵ_{cm} represents the error term, which consists of an idiosynchratic component and a term clustered on the department and month. As with any difference in difference design, the key underlying assumption for identification is that the control group serves as a valid counterfactual for the treatment group with parallel trends. Although we can not explicitly verify this assumption, we feel this threat is limited in this setting for several reasons. Because the strike was nationwide, and not just for the workers at oil refineries, any changes in response to the strike likely happened on a global scale that would have affected both the treatment and control groups. Moreover, the strike was a temporary condition, making it unlikely that workers relocated in search of new employment opportunities. Furthermore, because workers in France have health insurance regardless of employment status, there was unlikely to be a change in prenatal care consumption during the time of the strike. Figure 3.3 provides evidence to support the parallel trends assumption. Since there is little economic data available at such high temporal and spatial resolution, we plot the unemployment rate, which is available quarterly at the census tract, over time for the treatment and control groups. Although the unemployment rate is lower in census blocks with refineries, there is no trend difference between census blocks with refineries and their counterparts, supporting our contention that there are no differential trends across the two groups.

Figure 3.3: The distribution of unemployment by proximity to refineries



Note: Unemployment rates are available at the quarterly level for each census tract. 'Refinery areas' are census tracts where refineries are located, and 'non refinery areas' are census tracts without refineries.

3.4. Results

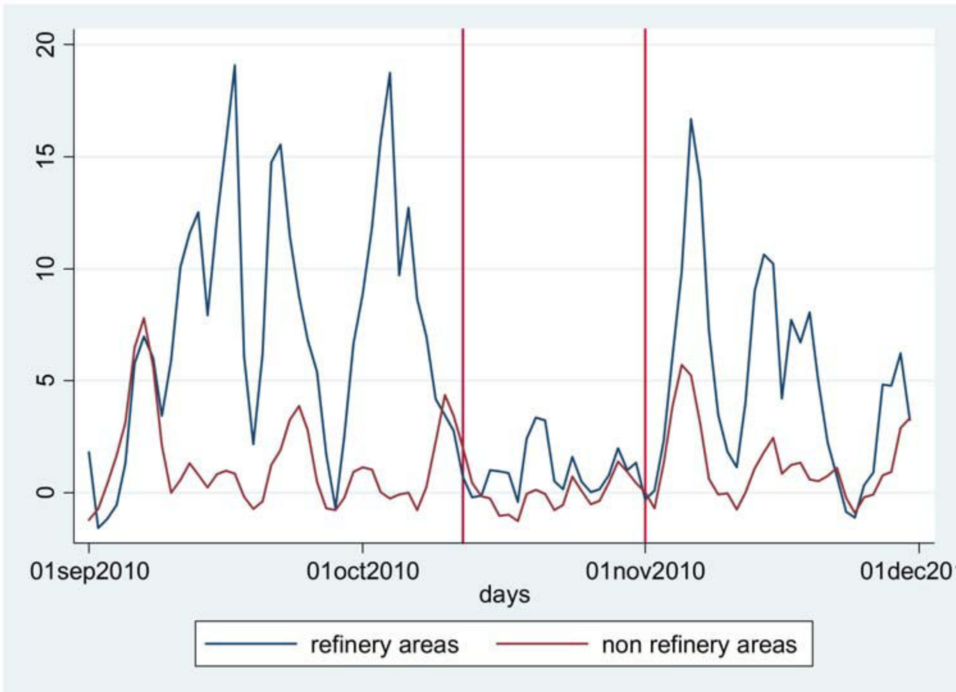
3.4.1. Refinery closures and pollution levels

We start by examining the effect of strikes on air pollution. The previously mentioned Figure 3.4 provides a daily graph of adjusted SO_2 pollution from September to December, 2010 for the treatment and control groups, with SO_2 adjusted by X_{cm} and σ_m. Prior to the strike, SO_2 levels are considerably higher in census tracts with refineries. However, during the strike, SO_2 dramatically falls in refinery areas to levels comparable to non-refinery areas. Immediately after the strike, SO_2 levels in refinery areas again exceed those of non-refinery areas. This visual display clearly demonstrates a strong, temporal effect of the strike on SO_2 levels.

Table 3.2 provides regression estimates of equation (3.1), which are largely analogous to this Figure. In order to gauge the extent of confounding, we successively add more time-varying controls, namely the weather variables and the unemployment rate. Consistent with Figure 3.4, the strike causes a statistically significant drop in SO_2 levels for areas close to refineries. SO2 levels drop during the strike by roughly 15 μg/m^3. Adding controls for weather (column 2) and unemployment (column 3) has no noticeable effect on our estimates.

The second and third panels explore the effect from different approaches for assigning pollution from monitors to census tracts. Limiting the sample to census tracts within 8 km of a monitor, shown in panel 2, leads to a slight increase in the effect of the strike on SO_2 levels. We see a much bigger increase, though still not

Figure 3.4: Adjusted SO_2 levels by proximity to refineries



Note: SO_2 levels are adjusted by weather variables, the local unemployment rate, and month and year dummy variables. The red lines indicate the approximate dates of the strike. 'Refinery areas' are census tracts where refineries are located, and 'non refinery areas' are census tracts without refineries.

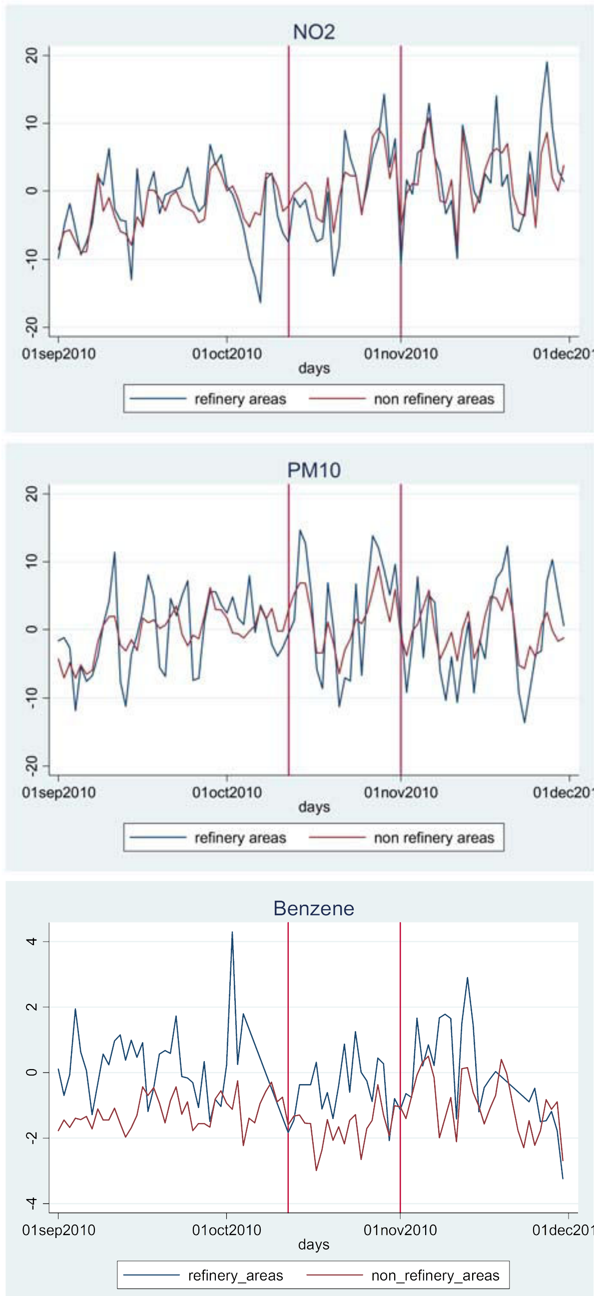
a statistically significant difference, when we limit to census tracts with 2 km of a monitor. This increase is consistent with a more precise measure of pollution from using a closer monitor. Overall, the results from Table 3.2, supporting the findings from Figure 3.4. Figure 3.5 presents the same plot as Figure 3.4 for three additional pollutants: NO_2, PM_{10}, and benzene.

Table 3.2: The effect of the strike on SO_2 levels

	1	2	3
A. All census tracts			
strike	-15.24*	-15.30*	-15.27*
	(8.796)	(8.799)	(8.772)
Observations	151,624	151,624	151,624
R-squared	0.758	0.758	0.758
B. Census tracts < 8km from monitor			
strike	-16.48*	-17.06*	-16.63*
	(9.020)	(9.065)	(8.713)
Observations	16,945	16,945	16,945
R-squared	0.757	0.758	0.758
C. Census tracts < 2km from monitor			
strike	-26.49**	-28.86**	-25.22**
	(11.23)	(11.30)	(10.79)
Observations	5,652	5,652	5,652
R-squared	0.756	0.757	0.757
weather		x	x
local economic conditions			x

Note: This table provides the coefficient estimates of the effect of strike on Sulfur Dioxide (SO_2). All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. The weather variables include average and maximum temperature, precipitation, minimum and maximum humidity, wind speed and direction. The unemployment rate is our measure of local economic conditions. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Figure 3.5: Adjusted NO_2, PM_{10}, and Benzene by area



Note: Pollution levels are adjusted by weather variables, the local unemployment rate, and month and year dummy variables. The red lines indicate the approximate dates of the strike.

While NO_2 and PM_{10} do not appear to change in response to the strike, Benzene shows a pattern consistent with being affected by the strike, though less stark than that for SO_2. While these patterns suggest SO_2 is the pollutant most affected by the strike, the possible relationship for other pollutants precludes us from conducting a proper instrumental variable (IV) analysis where we instrument SO_2 levels using the strike, though we cautiously provide IV estimates.

3.4.2. Refinery closures and birth outcomes

Given that we have found a relationship between the oil refinery strikes and pollution levels, we now turn our attention to the impacts of the strikes on health at birth. Tables 3.3 and 3.4 present results of the impact of exposure to the strikes anytime during pregnancy on birth weight and gestation, respectively. The top panel explores the effect on birth weight using the continuous measure and the low birth weight indicator, whereas the bottom focuses on gestational age and short gestation. Within each of the 4 dependent variables, we also explore sensitivity to controls as with the SO_2 results, as well as sensitivity to monitor-census tract distance assumptions. For birth weight, we find that birth weight increases by roughly 75 grams during the strike. This result is also insensitive to the addition of weather variables and unemployment. Compared to the mean birth weight of 3228 grams, this represents a 2.3 percent increase in birth weight. If we assume that the only pollutant affected by the refinery is SO_2, we can compute the effect of SO_2 on birth weight by dividing the effect of the strike on birth weight by the effect of the strike on SO_2 as shown in Table , akin to instrumental variables (IV). This procedure suggests that a 1 μg/m^3 decrease of SO_2 for one month increases birth weight by 5 grams, though we must interpret this with caution because, as noted above, the refineries may have affected other pollutants, such as benzene, which would make IV valid.

Table 3.3: The effect of the strike over the entire pregnancy on birth weight

	1	2	3	4	5	6
		birth weight (g)			birth weight < 2500 g	
A. All census tracts						
strike	73.61*	76.47*	76.44*	-0.020*	-0.021*	-0.021*
	(44.61)	(44.75)	(44.73)	(0.011)	(0.011)	(0.011)
Observations	121,157	121,157	121,157	121,157	121,157	121,157
R-squared	0.053	0.053	0.053	0.066	0.066	0.066
B. Census tracts < 8km from monitor						
strike	71.87	74.87*	74.03*	-0.019*	-0.020*	-0.019*
	(44.50)	(44.91)	(44.83)	(0.011)	(0.012)	(0.012)
Observations	14,169	14,169	14,169	14,169	14,169	14,169
R-squared	0.043	0.044	0.045	0.066	0.067	0.067
C. Census tracts < 2km from monitor						
strike	92.38*	99.43**	99.03**	-0.025*	-0.026*	-0.026*
	(47.21)	(48.41)	(48.41)	(0.014)	(0.014)	(0.014)
Observations	4,962	4,962	4,962	4,962	4,962	4,962
R-squared	0.055	0.059	0.060	0.049	0.054	0.054
weather		x	x		x	x
local economic conditions			x			x

Note: This table provides the coefficient estimates of the effect of exposure to the strike at any time during pregnancy on birth weight. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

Using an indicator for low birth weight, we find that the strike lowered this rate by roughly 2 percentage points, which is also statistically significant and robust to additional controls. When we limit the distance from pollution monitor to the census tract to 8 km, our estimates change minimally, as with the SO_2 results. Limiting to 2 km leads to a larger improvement in birth weight, though the difference is again not statistically significant.

For gestational age, we find similar qualitative results. Using all census tracts,

Table 3.4: The effect of the strike over the entire pregnancy on gestation

	1	2	3	4	5	6
		gestation (wks)			gestation < 37 wks	
A. All census tracts						
strike	0.361*	0.382*	0.383*	-0.091***	-0.094***	-0.094***
	(0.194)	(0.196)	(0.195)	(0.029)	(0.029)	(0.029)
Observations	90,134	90,134	90,134	90,134	90,134	90,134
R-squared	0.071	0.071	0.071	0.075	0.075	0.075
B. Census tracts < 8km from monitor						
strike	0.366*	0.373*	0.373*	-0.088***	-0.087***	-0.087***
	(0.196)	(0.197)	(0.197)	(0.030)	(0.030)	(0.030)
Observations	10,761	10,761	10,761	10,761	10,761	10,761
R-squared	0.081	0.083	0.083	0.087	0.089	0.089
C. Census tracts < 2km from monitor						
strike	0.375	0.407*	0.400*	-0.062*	-0.066*	-0.065*
	(0.243)	(0.242)	(0.241)	(0.037)	(0.036)	(0.036)
Observations	3,849	3,849	3,849	3,849	3,849	3,849
R-squared	0.111	0.120	0.120	0.111	0.121	0.121
weather		x	x		x	x
local economic conditions			x			x

Note: This table provides the coefficient estimates of the effect of exposure to the strike at any time during pregnancy on gestation. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

regardless of distance to a pollution monitor, we find the strike increased gestational age by roughly 0.37 weeks, or 2.5 days, which is a 1% change from the baseline mean. This yields an IV estimate of a 1 μg/m^3 decrease of SO_2 for one month increases gestational length by 0.18 days. The strike reduces the probability of short gestation by .08. These results are again insensitive to additional controls. While the results do not become larger when limiting to a shorter distance from the census tract to the pollution monitor, the differences are again not statistically significant. To compare the estimates for birth weight and gestation, we perform the following calculation. Since the fetus gains about 200 grams in weight per week in the final month of pregnancy (Cunningham *et al.*, 2005), the 0.37 week increase in gestation translates into an extra 74 grams in weight, which is nearly identical to our estimate on the impact on birth weight. Therefore, it appears that the reduction in birth weight is solely due to shorter gestation, rather than growth retardation. Since the strike only lasted for less than one month, as previously mentioned one of the advantages of our study is the ability to more precisely isolate the effects by trimester. Table 3.5 presents results by including exposure to the strike by trimester. We focus solely on census tracts less than 8 km from a monitor and with the meteorological and economic covariates included, though results are robust to different assumptions regarding these choices.

Table 3.5: The effect of the strike on birth weight and gestational age by trimester of pregnancy, census tracts within 8 km of pollution monitor

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
strike - 3rd trimester	151.2***	-0.024**	0.847***	-0.110***
	(50.15)	(0.012)	(0.226)	(0.031)
strike - 2nd trimester	10.63	-0.019	0.133	-0.082***
	(66.14)	(0.012)	(0.300)	(0.030)
strike - 1st trimester	60.02	-0.015	0.138	-0.069**
	(78.78)	(0.012)	(0.250)	(0.033)
Observations	14,169	14,169	10,761	10,761
R-squared	0.045	0.067	0.083	0.089
weather		x	x	
local economic conditions			x	

Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation by trimester of pregnancy when the distance from the census tract to the pollution monitor is less than eight kilometers. All specifications include census tract fixed effects, year and month dummy variables, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

We find that almost all of the effects from pollution are due to exposure during the third trimester. Birth weight increases by roughly 150 grams when the strike occurred during the third trimester, which represents a 4.6 percent increase. The effects from the first and, in particular, second trimesters are much smaller and not statistically significant. Turning to the incidence of low birth weight, we find reasonably similar effects across the trimesters, but the third is the largest (and comparable to the estimate for the overall pregnancy) and the only one that is statistically significant. For gestational age, we also find that exposure to the strike in the third trimester has the biggest effect: it increases gestational age by roughly 0.85 weeks, a roughly 2.2 percent increase. This longer gestation translates into roughly 170 grams, which again explains all of the estimated effect on birth weight from third trimester exposure. The effects in the first and second trimester are again much smaller and not statistically significant. Turning to the incidence of short gestation, we again find the third trimester has the biggest effect, but the first and second also appear significantly related to short gestation.

Table 3.6: Estimates using alternative measure of strike exposure, census tracts within 8 km of pollution monitor

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
	A. Entire pregnancy			
strike	73.50	-0.0193*	0.347*	-0.0839***
	(44.83)	(0.0114)	(0.198)	(0.0298)
Observations	14,169	14,169	10,769	10,769
R-squared	0.045	0.067	0.083	0.089
	By trimester			
strike - 3rd trimester	148.1***	-0.0224*	0.815***	-0.106***
	(49.92)	(0.0118)	(0.225)	(0.0306)
strike - 2nd trimester	8.323	-0.0184	0.110	-0.0810***
	(66.62)	(0.0121)	(0.307)	(0.0298)
strike - 1st trimester	63.69	-0.0171	0.111	-0.0643*
	(79.75)	(0.0121)	(0.249)	(0.0336)
Observations	14,169	14,169	10,769	10,769
R-squared	0.045	0.067	0.083	0.089

Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation assuming all births occurred on the 15th of the month (as opposed to 1st). All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

As previously mentioned, we do not know the exact date of birth of the child, only the month of discharge from the hospital. In Table 3.6, we present results assuming the date of child's birth is on the 15th of the month instead of the 1st, again focusing solely on the census tracts within 8 km of a monitor. Our results from this specification are virtually identical to the main results, suggesting the lack of knowledge about the exact birth date is not hindering inference. Since pollution and other environmental confounders often show strong seasonal patterns, we want to ensure that our results are not driven by this phenomenon. To assess this, we present estimates from a falsification test where we assign the date of the strike to have occurred on October, 2009, a year before the actual strike occurred. Shown in Table 3.7, we find that the placebo strike is neither associated with SO_2 levels or any of the birth outcome measures. Of the 17 coefficients shown, only 1 is statistically significant (at the 10% level), which is almost exactly what we expect given the chance of a Type I error.

Table 3.7: Effect of placebo strike in October, 2009, census tracts within 8 km of pollution monitor

	1 SO_2	2 birth weight (g)	3 birth weight < 2500 g	4 gestation (wks)	5 gestation < 37 wks
A. Entire pregnancy					
strike	0.112	44.02	0.023	-0.149	0.050
	(1.608)	(68.54)	(0.038)	(0.379)	(0.071)
Observations	16,945	14,169	14,169	10,761	10,761
R-squared	0.619	0.045	0.067	0.083	0.089
B. By trimester					
strike - 3rd trimester		69.28	-0.013	0.044	0.040
		(95.03)	(0.009)	(0.350)	(0.115)
strike - 2nd trimester		65.82	-0.018*	0.186	0.049
		(96.34)	(0.010)	(0.437)	(0.112)
strike - 1st trimester		-3.07	0.101	-0.676	0.060
		(151.70)	(0.107)	(0.903)	(0.113)
Observations		14,169	14,169	10,761	10,761
R-squared		0.045	0.068	0.083	0.089

Note: This table provides the coefficient estimates of the effect of a placebo strike occurring October, 2009 on SO_2, birth weight and gestation. All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

The distance to the refinery should also be an important factor of the effect on health. In table 3.8, we modify the treated group by including census tracts within a certain distance to the closest refinery. As shown in the first block, when the treated group corresponds to census tracts within 2 km of the refinery, the positive effect of the strike on birth weight and gestational age is strong and highly significative. Estimates for birth weight are reducing but are not significative anymore when we increase the treated group by including census tracts within 5 km of the refinery. However, gestational age estimates remains significative with a lower coefficient. Increasing the treated group to census tracts within 10 km of the refinery removes significativity. Results are in line with the intuition: the effect on health tends to disappear further away we are from the source of pollution.

Table 3.8: Estimates using alternative distance between census tracts and refineries

	1 birth weight (g)	2 birth weight < 2500 g	3 gestation (wks)	4 gestation < 37 wks
A. treated group: census tracts around 2 km				
strike	87.22***	-0.0223**	0.338**	-0.0636**
	(27.29)	(0.0106)	(0.154)	(0.0248)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065
B. treated group: census tracts around 5 km				
strike	20.22	-0.00842***	0.0879	-0.0240
	(17.05)	(0.00313)	(0.0956)	(0.0315)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065
C. treated group: census tracts around 10 km				
strike	-19.71	0.00471	-0.00286	-0.00709
	(14.91)	(0.00560)	(0.0615)	(0.0158)
Observations	121,157	121,157	90,134	90,134
R-squared	0.044	0.056	0.062	0.065

Note: This table provides the coefficient estimates of the effect of strike on birth weight and gestation using alternative distance for the treated group between refinery and census tract . All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1.

The way we assign pollution to census tract is an important factor. It is indeed possible that a census track is localized close to monitoring stations situated in a neighborhood department. To take into account department border effect, we take now the weighted average of pollution measurements from both all monitors within the department and all monitors within a neighborhood department using distance between the census tract centroid and the monitoring station as weights (Table 3.9). Results are similar to the previous one but the size of the effect is higher. SO_2 significantly decrease by about 30 μg/m^3 which has a stronger effect on birth weight and gestational age. We again find a significant effect in the third trimester of pregnancy. These results suggest the effect may be stronger closer we are from the monitor.

Table 3.9: Estimates using alternative measure of exposure to air pollution

	1 SO_2	2 birth weight (g)	3 birth weight < 2500 g	4 gestation (wks)	5 gestation < 37 wks
A. Entire pregnancy					
strike	-32.15***	91.75*	-0.0268*	0.365*	-0.0923***
	(49.95)	(0.0146)	(0.195)	(0.0297)	
Observation	151,624	121,157	121,157	90,134	90,134
R-squared	0.811	0.063	0.077	0.072	0.076
B. By trimester					
strike - 3rd trimester		156.3***	-0.0274*	0.810***	-0.0996***
		(56.70)	(0.0148)	(0.219)	(0.0301)
strike - 2nd trimester		20.66	-0.0275*	0.0815	-0.0900***
		(67.91)	(0.0146)	(0.307)	(0.0293)
strike - 1st trimester		98.26	-0.0255*	0.204	-0.0871***
		(83.25)	(0.0151)	(0.243)	(0.0313)
Observations		121,157	121,157	90,134	90,134
R-squared		0.053	0.067	0.083	0.089

Note: This table provides the coefficient estimates of the effect of strike on SO_2, birth weight and gestation using alternatives measure of exposure. All specifications include census tract fixed effects, year and month dummy variables, weather, and local economic conditions, with standard errors clustered at the month and department level in parenthesis. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

3.5. Conclusion

The goal of this paper was to examine an externality from energy production, focusing on health impacts as measured by birth outcomes. To account for the endogeneity of pollution exposure, we exploit the oil refinery strike that occurred in October 2010, which led to a sharp, temporary reduction in SO_2 in areas close to the refineries. This reduction led to a robust increase in birth weight and gestation of infants, particularly those who were exposed during their third trimester of pregnancy. To gauge the magnitude of these estimates, we perform the following illustrative calculations, similar to Currie *et al.* (2009). We value the improvements in birth weight by computing the percentage change in birth weight from the change in pollution in October, 2010 by dividing the estimated impact of third-trimester SO_2 on birth weight from Table 3.5 (140) by the mean birth weight in our sample (3220) from table 3.1. We multiply this by the estimated elasticity between birth weight and earnings of 0.1 from Black *et al.* (2007) to obtain the percentage change in earnings during the month of strikes. We then multiply this by the average gross annual earnings of all full time workers (33,168 euros) from the Directorate for Research, Studies, and Statistics in 2010 in France. Finally, we multiply by the total number of births in 2010 (832,799) to get the change in earnings per year. This gives an estimated increase in nationwide earnings of 120 million euros. Assuming a 40 year working career with 3 percent annual rise in earnings and a 6 percent discount rate, this amounts to 2.933 billion euros per cohort. If we attribute all of the estimated 15 unit decline in SO_2 to the strike, this implies that a 1 unit decrease in

SO_2 increases future earnings of a given birth cohort by 196 million euros per year.[12] While only meant to be illustrative, these estimates suggest that the externalities from oil production that accrue to newborns alone are potentially quite sizeable and should be an important part of policy discussions surrounding the production of energy. Further research regarding this issue may also look at social inequalities in health. Combining an exogenous environmental shock with data on infant health and mother's social characteristics would shed light on disparities issues.

12. Clearly, these estimates understate the full benefits from a decrease in SO_2 because they only capture the earnings impacts for a birth cohort and only capture the effects on births. A 3% (4%) discount rate would yield an earnings increase of 328 (272) million euros per 1 unit change in SO_2.

Chapter 4

The Price of Pollution and Health: A Hedonic Approach

This paper examines the impact of a reduction in sulfur dioxide concentration (SO_2) in France on both health outcomes and property prices, at a municipality level, from 2008 to 2011. The paper aims at putting in light people's willingness to pay for perceived differences in environmental attributes and to avert the real cost in terms of health they are exposed to. To do so, I apply a hedonic price analysis and a damage function method to the recent closure affecting a oil refinery in the north of France, in September 2009, as a natural experiment. This contribution shows, first, that the permanent closure of the plant leads to a reduction in SO_2 concentration. Although we find an objective and positive effect on health, people perception, derived from the hedonic approach, does not always respond positively to this air quality improvement. The estimates suggest that the environmental effect is smaller in relative terms for buyers of cheaper flats than the economic effect.

4.1. Introduction

The decision to live next to a toxic site may be associated to a wide number of concerns. As Portney (1981) suggested it may be possible to draw inferences about individuals' valuation of risk by combining estimates of the effect of air pollution on both property values and human health risks. This paper aims to evaluate the consequences of atmospheric pollution after the closure of a toxic site.

Rosen (1974) hedonic model shows that the willingness to pay for a change in natural resource can be inferred by the explicit price of a property. Chay & Greenstone (2005) find that the elasticity of housing values with respect to particulates concentrations ranges from -.20 to -.35. Greenstone & Gallagher (2008) look at areas chosen for the Superfund-sponsored cleanups of hazardous waste sites program compared to their counterparts. They find that Superfund cleanups are associated with economically small and statistically insignificant changes in residential property values, property rental rates, housing supply, total population, and types of individuals living near the sites. The estimates from Hanna (2007) suggest that being a mile closer to a polluting manufacturing plant reduces house values by 1.9%. Besides, the literature focuses also on the relationship between health and housing prices. The superfund cleanups of hazardous waste sites has also been used to shed light on its impact on health status at birth (Currie *et al.*, 2011). Davis (2004) measures the effect of health risk on housing values by exploiting an isolated county in Nevada where residents have experienced a severe increase in pediatric leukemia. The estimated MWTP to avoid pediatric leukemia risk is used to calculate the value of a statistical case of pediatric leukemia. Most recently, Currie *et al.* (2013) looked at the housing market and health impacts of 1600 openings and closures of industrial plants that emit toxic pollutant. The paper shows that housing values within one mile decrease by 1.5 percent when plants open, and increase by 1.5 percent when plants close.

While the hedonic price analysis focuses on the perception people have about a reduction in health risk or in air pollution and on their behaviour in reaction, the damage function is an objective measure of the impact of air pollution on health.

The hedonic price analysis (HPA), by using the housing market to estimate the economic benefits of air quality, only captures people's willingness to pay for perceived differences in environmental attributes, and their direct consequences. However, the individuals' perception of health risk may differ from the real risk population living near toxic sites is exposed to. Thus, the article aims at analyzing the differences between both methods. Though a wide literature deals with the link between pollution and hedonic prices, existing studies do not simultaneously adress both the willingness to pay to avoid negative externalities reflected in housing price differential and the real cost in terms of health risks people are exposed to. In this perspective, it may be tempting to directly compare the willingness to pay to avoid negative externalities reflected in housing price differential and the real cost in terms of health people are exposed to. However, on the one hand, the damage function is likely to underestimate the total health benefits in this study because it omits the effet of pollution on mortality. A large number of studies focusing on the economic value of reducing SO_2 and related pollutant find that a high part of the quantified health benefits are associated with mortality reduction. More than 80% of monetized benefits were attributed to reductions in premature mortality (Krupnick *et al.*, 2002). The order of magnitude of health benefits are larger than the ones we can report in this study as we only focus on respiratory outcomes.

In addition, disutility from degraded health is only one part of people's perception of air pollution. Not only may the health risk due to air pollution be a source of disutility but also environmental amenities: the view of a factory chimney, the smoke or possible odor may have an impact on the perception of the population living near a refinery. On the other hand, the hedonic prices analysis ignores the effect of air pollution on the property market as a whole and its long term effects. Because a number of constraints prevent comparison between both methods, I chose to put in parallel estimation results of environmental benefits from both methods to see whether diagnoses converge. In other words, this global study on air pollution, health outcomes and housing prices analyzes to what extent the hedonic price analysis (HPA) captures people's willingness to pay for perceived differences in environmental attributes (perceived health risk from pollution and perceived environmental amenities) and how it may differ from the objective and biological environmental health risk.

The analysis focuses on Dunkirk, a French municipality in the Nord-Pas de Calais region, and not only the area of Dunkirk, in France where residents have recently experienced a refinery closure. Pollution, health outcomes and housing prices are compared before and after the closure with the nearby municipalities, within 50 kilometers of Dunkirk, with at least one monitoring station, acting as a control group.

People living near the refinery may actually be concerned differently by the refinery closure. In fact, in the case of the refinery closure in Dunkirk, two types of effects occured: economic and environmental. First, refinery closure produced an economic shock in term of employment and economic activity in the surrounding of Dunkirk. As the population became unemployed, the economic dynamism of the area fall. Such activity slowdown decreased the wealth of people living in the area of the refinery. Economic activity can have a direct impact on prices and may bias the estimate. Nevertheless, this economic shock affected the entire Nord-Pas de Calais region in terms of economic outcomes due to direct and indirect employment consequences of the refinery closure.

Second, the environmental shock corresponds to a decrease in pollution concentration in the area of Dunkirk following the refinery closure. Thus, I assume that the closure of the refinery may have reduced health risks from air pollution, likely increased the aesthetic value of proximity to the site and finally, enhanced the value of neighborhood properties.

I first show that the closure of the refinery leads to a reduction in air pollution; I use this reduction in air pollution to infer the impact it has on hospital respiratory outcomes. Keeping in mind the positive effect of the refinery closure on health, I show, in parallel, that the effects of the refinery closure on property prices are not always the ones we expect. This valuation may be of use in determining if the willingness to pay for the perceived pollution reduction in different segments is in line with the economic benefits of an improvement in air pollution concentration.

4.2. Pollution, health and refinery closure

4.2.1. SO_2 pollution and health

This paper focuses on sulfur dioxide (SO_2), one of the major pollutant emitted by oil refineries and the main industrial pollutant. Sulfur dioxide (SO_2) is one of a group of highly reactive gasses known as oxides of sulfur (SO_x). The largest sources of SO_2 emissions are from fossil fuel combustion at power plants and other industrial facilities (Environmental Protection Agency (EPA), 2011). SO_2 is a colorless gas with a very strong smell. SO_2 is subject to transformation in the atmosphere and can react with other compounds to form small particles. These particles go deeply to lungs and can cause or aggravate respiratory diseases, such as emphysema and bronchitis.

Subjects exposed to SO_2 showed decreased lung functioning for children and increased respiratory symptoms for adults (World Health Organization (WHO), 2011), asthma crisis and ocular rash (Lecoq *et al.*, 2009). Inflammation of the respiratory tract causes coughing, mucus secretion, aggravation of asthma and chronic bronchitis and makes people more prone to infections of the respiratory tract (World Health Organization (WHO), 2011). The effects seem stronger for high levels of exposure and people with asthma are more sensitive to SO_2. The number of hospital admissions for cardiopathy and mortality increases on days with a high SO_2 air concentration (Finkelstein *et al.*, 2003). Human clinical studies consistently demonstrate respiratory morbidity among exercising asthmatics following peak exposures (5-10 min) to SO_2 concentrations equals 0.4 ppm, with respiratory effects occurring at concentrations as low as 0.2 ppm for some asthmatics (World Health Organization (WHO), 2005).

4.2.2. Pollution and the refinery closure

Refineries are responsible for 20 % of SO_2 release in France (Soleille, 2004). Oil refineries convert crude oil to everyday product like gasoline, kerosene, liquefied petroleum. Crude oil and coal contain a relatively high quantity of sulfur. SO_2 is created when crude oil or coal is heated at the refinery to produce fuel. Thus, the refining process releases a large number of chemicals such as benzene, chromium and sulfur acid into the atmosphere. Therefore, refineries are considered upper tier SEVESO sites for most of their activities. In Europe, the "Seveso" directive applies to around 10,000 industrial establishments where dangerous substances are used or stored in large quantities, mainly in the chemicals, petrochemicals, storage, and metal refining sectors. The Seveso Directive compels Member States to ensure that operators have a policy in place to prevent major accidents.

The Flandres refinery, close to Dunkirk in northern France, is part of SEVESO sites. The refinery produces liquefied gases (propane and butane), fuel for airplanes and automobiles (gasoline and diesel), domestic and industrial fuel and biofuels. Its refining capacity is up to 7.8 million tons/annum. The refinery employs nearly 370 employees on average, which represents 0,4% of the global employment in Dunkirk. The refinery also annually works with 775 establishments which generates 87 million euros of turnover, among 275 establishments localized in Nord-Pas de Calais, and account for 44,1 million of euros. Most of the employees live near their workplace: two thirds of the employees of the refinery live in Dunkirk. Most of them work full-time, under a permanent employment contract and are labor workers. One quarter of the refinery workers belongs to intermediate profession (INSEE, 2010). In september 2009, the production of the refinery is shutted down due to poor demand and margins. Given the poor outlook, French oil giant Total announced definitely the closure of its refinery in 2010. I reasonably believe that the closure had an impact on all the Nord-Pas de Calais region affecting not only refinery workers but also a large range of subcontractors from all the Nord-Pas de Calais region.

4.2.3. The incidence of the refinery closure on housing prices

Hedonic model The hedonic price model, derived mostly from Lancaster (1966) consumer theory and Rosen's (1974) model, assumes that a differentiated good can be described by a vector of its characteristics. The hedonic approach to evaluation aims to estimate the economic value of a good using implicit price of the product attributes. In the case of a property, these characteristics may include structural attributes (e.g., number of bedrooms), neighborhood public services (e.g., local school quality), and local environmental amenities (e.g., presence of a toxic site). People have the opportunity to select the combination of features they prefer, given their income. In this context, areas with elevated health risks such as Dunkirk must have lower housing prices to attract potential homeowners.

However, valuations derived from hedonic prices functions must be interpreted carefully. The literature emphasizes a wide number of critics of the HPAs.

Under perfect information, the price differential associated with proximity to hazardous sites reflects both individuals' valuations of the greater health risk and any effect on neighborhood aesthetics (Greenstone & Gallagher, 2008). Although a biological health risk may exist, it is not clear the extent to which individuals are fully and correctly informed about the health impacts of air pollution. Although individuals may be aware about air pollution, it is less likely that they correctly incorporate this risk into their pricing decisions for housing. Imperfect information suggests that the hedonic approach underestimates the true health cost of air pollution whereas damage-function may tend to overestimate the health costs because mortality may be too high (Delucchi *et al.*, 2002). Cropper (2000) notes that it is more likely that the property values used in HPAs capture all of the aesthetic benefits, but only capture a portion of the health benefits. Zabel & Kiel (2000) emphasize it is still unclear how individuals process air quality information when determining their willingness-to-pay for housing.

Another controversial issue is that of market segmentation. Feitelson *et al.* (1996) noted that in theory, hedonic price studies do not require the segmentation of housing markets if all attributes are adequately taken into account. However, in practice, several types of market segmentation are likely to exist in most markets. This is because housing markets are not uniform (Adair *et al.*, 1994),(Fletcher *et al.*, 2000). Hence, it is unrealistic to treat the housing market in any geographical location as a single entity. For instance, house and flat should be considered separately due to their specificity and differences in price evolution. Unfortunately, the definition, composition, and structure of sub-markets have not been given much attention in the hedonic-price literature, although it is an important empirical issue.

A remaining issue frequently associated with the hedonic price model is pollution endogeneity. HPAs consider that pollution is an exogenous variable in the regression of housing prices. This is not always correct. Industrial facilities, sources of pollution, are probably located in areas with specific characteristics such as low population and relatively low housing prices. On the one hand, employees from an industrial company are willing to live close to their place of work so as to limit their everyday transport. On the other hand, atmospheric pollution reduces air quality and the attractivity to live nearby pollution sources. In this context, property prices should decrease. Bajari *et al.* (2012) suggest that ignoring bias from time-varying correlated unobservables considerably understates the benefits of a pollution reduction policy. In fact, the presence of omitted variables may be correlated with the pollutants. Confounding factors, such as the opening of new businesses, may evolve over time in conjunction with worsening air pollution. It appears that failing to control for omitted attributes or controlling for time-invariant unobserved attributes with fixed effects only leads to the wrong sign on the estimate of the potential benefits of a pollution reduction policy. When differences between locations are imperfectly measured and covary with health risk and housing prices it becomes difficult to disentangle the price effects of health risks from the price effects of other locational amenities (Davis, 2004).

As developed by Chay & Greenstone (2005), differences in terms of pollution preferences may also lead to autoselection bias. Households may sort themselves to locations endowed with amenities that match their preferences. For instance, people with low preferences for air quality may sort themselves into location with a high level of air quality. The subsample studied may not be representative of the whole population. In this case, hedonic estimation only reflects marginal prices of air quality for a part of the population who do not value air pollution. The value of the marginal price of air pollution will be underestimated.

A last issue issue frequently associated with the hedonic price model is the misspecification of variables. Over-specification, inclusion of the irrelevant variable, gives estimated independent variables that are both unbiased and consistent, but inefficient whereas under-specification results in estimated coefficients that are both biased and inconsistent. According to Butler (1982), since all estimates of hedonic price models are to some extent misspecified, models that use a small number of key variables generally would be suitable.

To mitigate these problems and to infer the impact of a reduction in air pollution on housing prices and health status, I use a quasi-experimental approach. The analysis focuses on Dunkirk, a French municipality in the Nord-Pas de Calais region in France where residents have recently experienced a refinery closure. Besides, the conditions of supply and demand are relatively similar in the Nord-Pas de Calais property market such that I can expect a similar set of implicit prices in Dunkirk and in its surrounding analyzing flats and housing separately. People living in Dunkirk may have a high-risk behavior leading to a low health status. The stop in the refining process allows to well measure the effect of the pollution on population exposed to pollution emitted by the refinery compared to the population living far from the refinery and not exposed to its pollution. The use of variation in pollution and health risk over time is of particular interest to control for unobserved differences across locations. Using a really rich and exhaustive dataset about property transactions, I am also able to shed light on population sorting. Kuminoff *et al.* (2010) suggest that large gains in accuracy can be realized by moving from the standard linear specifications for the price function to a more flexible framework that would use a combination of spatial fixed effects, quasi-experimental identification, and temporal controls for housing market adjustment. Taking these elements into account, I analyze in this study the link between the reduction in air pollution, due to the refinery closure, hospital respiratory outcomes and the property value.

Theoretical basis Before entering in more details to empirical results, I develop, in this section, a theoretical model to compare willingness to pay across segments. I adapt a model from Brookshire *et al.* (1982). The starting point for utility theoretic analysis is the household's utility maximization problem. The household's utility function is denoted $U(X; P; \vec{Z})$. X represents a private composite, P is the level of air pollution and $\vec{Z}$ represents the vector of the housing attributes, including the type of property (house or flat). The variable P represents an environmental quality variable namely the level of pollution and is assumed to be a public bad not chosen by households. It enters the preference structure such that the marginal utility of pollution is decreasing $\partial u/\partial P < 0$. On the opposite, the utility is an increasing function of consumption $\partial u/\partial X > 0$.

Each household maximizes utility which is subject to the full income budget constraint:

$$Y - CX - R(P, \vec{Z}) = 0$$

where Y is the household income, C is the unit cost or the price of the commodity and R is the rent gradient as defined by Rosen. As usual in this literature, I assume a decline in rents when the distance from the polluted site decreases, or equivalently, a decrease in rents when the pollution increases $\partial R/\partial P < 0$. The first order conditions for choice of P and X imply:

$$C\frac{U_P}{U_X} = R_P(P, \vec{Z}); C\frac{U_{\vec{Z}}}{U_X} = R_{\vec{Z}}(P, \vec{Z})$$

Thus a hedonic rent gradient is defined for pollution P and different types of property as well.

Let us assume two households A and B with identical preferences but different income ($Y^A > Y^B$). As illustrated in the graph 4.1, if the initial level of pollution is given by P_0, household A, located at P_0 , has an indifference curve I^A while household B, has an indifference curve I^B, also located at P_0. Household A, with income Y^A, would then face a rent gradient like that shown in figure 4.1 defined by $R(P)$ [1] and choose point a and household B, with income Y^B, would then face a rent gradient defined by $R(P)$ and choose point b. Therefore, the graph shows that households with different levels of income, after a change in pollution from P_0 to P_1, may face the same rent gradients over pollution and that their absolute WTP should be equal. Nevertheless, perception about the level of pollution can be misperceived. If a poorly informed household underestimates the change in pollution from P_0 to P_1' due to a lack of awareness, the rent differential from the well informed household will exceed the rent differential from the poorly informed household. The graph is drawn under the assumption that income and information (or the ability to use information) are correlated because of the usual evidence about the link between education and income (Griliches & Mason, 1972). But his informational failure could have different consequences and the final effect is ambiguous on the difference between the rent differentials that affect both households. The poorly informed household could underestimate the initial level of pollution P_0, or underestimate or also over-

1. Since we assume that the households share the same preferences, there is no reason for them to choose different housing attributes for a given level of pollution. This is the reason why we omit $\vec{Z}$ here.

estimate the difference P_0 - P_1. It is thus clear that such an informational failure introduces some noise in the differential rent measures that can explain some discrepancies between classes of households. As shown by Brookshire *et al.* (1982), the graph emphasizes again that the rent differential always underestimates the WTP of each type of households. Moreover, in case of imperfect information about the real level of pollution (or its health effects), even the absolute willingness to pay may differ. Thus, information plays an important role in the perception of pollution and may lead to different rent gradients.

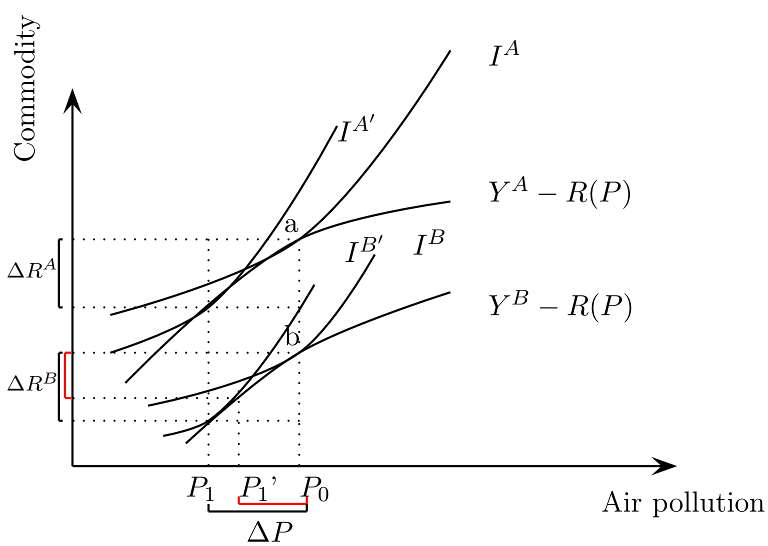


Figure 4.1: Rent evolution with a lack of information

One should observe what is happening now if we consider different preferences for the agents, allowing to distinguish different segments such as houses and flats. As illustrated in the graph 4.2, household A, located at P_0, with an indifference curve I^A, buys a house $\vec{Z}^A$ while household B, with an indifference curve I^B, also located atP_0 buys a flat $\vec{Z}^B$. Household A, with income Y^A, would then face a rent gradient like that shown in figure 4.2 defined by $R(P, \vec{Z}^A)$ and choose point a but household B, with income Y^B, would then face a rent gradient defined by $R(P, \vec{Z}^B)$ and choose point b. Therefore, households with different type of property, after a change in pollution from P_0 to P_1, may face different rent gradients over pollution as emphasized in the graph 4.2. The change in the rent gradient of a house ΔR^A exceeds the change in the rent gradient of a flat ΔR^B. Thus, people living near the refinery in a flat or a house may actually be concerned differently by the refinery closure.

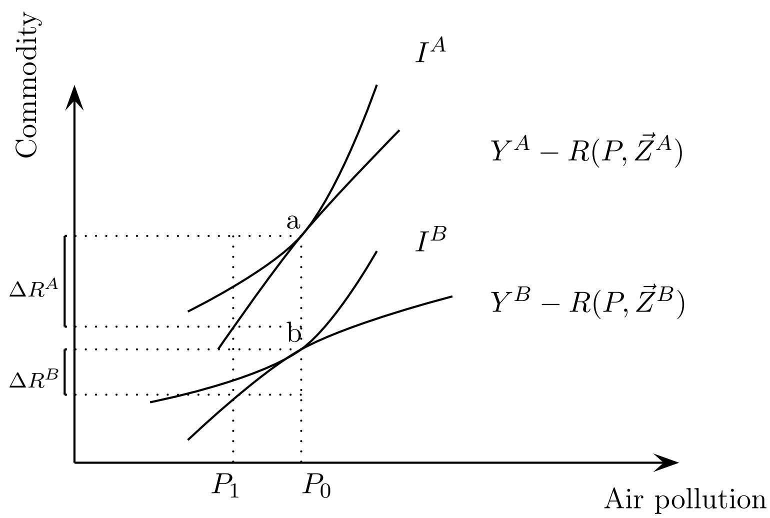


Figure 4.2: Rent evolution with respect to segments

If we want to address the role of different spending constraints, let us assume that each household has the same minimal level X_0 of commodity consumption (to satisfy the primary needs of households besides housing, like energy, water or minimal food consumption).

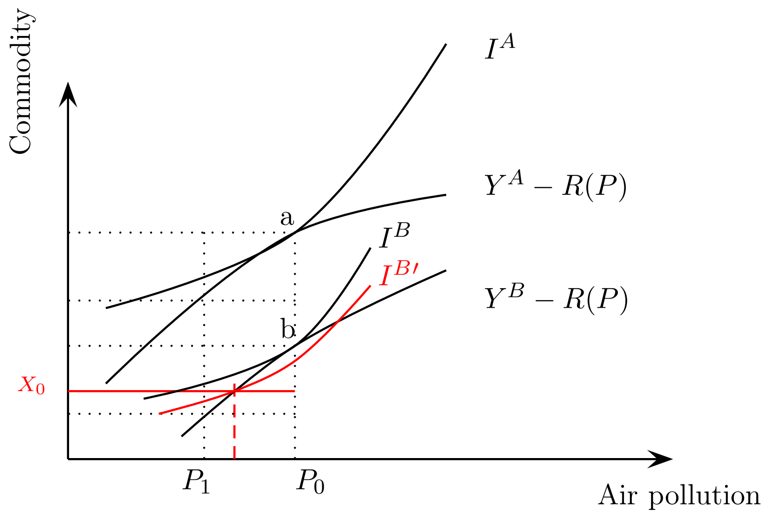


Figure 4.3: The role of different spending constraints

In that case, with X_0 binding, the household cannot freely maximize their utility and have to "choose" a level of rent corresponding to a higher level of pollution than the real one (Figure 4.3). Their WTP is constrained by their budget constraint.

4.3. Dataset presentation

4.3.1. Pollution data

Air quality is monitored throughout France (mainland and overseas departments) by 38 approved air quality monitoring associations (AASQA). The French monitoring station system counts approximately 700 measurement stations equipped with automatic instruments and nearly 400 experts implement this monitoring system. I focus on sulfur dioxide (SO_2) concentration in the Nord-Pas de Calais region around 50 kilometers from Dunkirk which represents two departments, geographical level below the regional level. I obtain daily measures of ambient air pollution concentrations in microgram per cubic meter ($\mu g/m^3$) for all air quality monitors in France for 2008-2011 from the Ministry for Ecology, sustainable development and spatial planning (ADEME) database and more recently from the national institute of industrial environment and risks (INERIS).

Table 4.1 presents the summary statistics of all the variables. Monthly pollution concentration data are presented in panel A of the summary statistics where I present a measure of expected exposure to SO_2 after having dropped stations that do not exist for the entire period: from 2008 to 2011. In addition I remove monitoring stations that do not measure SO_2. Note also that only some municipalities dispose of a monitor. These 2 departments represent 238 municipalities and 16 air pollution monitoring stations.[2] The distribution of monitoring stations throughout France is represented in figure 4.4 with a marquee which represents the area of the study, 50 kilometers around Dunkirk, in the Nord-Pas de Calais.

The summary statistics indicates that the monthly SO_2 concentration decreases after the refinery closure. The level is quite low over the period due to the monthly aggregation from daily data. Sulfur dioxide concentration after the refinery closure in Dunkirk decreased from 12.65 to 6.58 $\mu g/m^3$ whereas SO_2 concentration in the other municipalities, decreased from 3.12 to 2.21 $\mu g/m^3$ in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is 5.16 $\mu g/m^3$. I have also performed an independent samples t-test to compare the means of a normally distributed interval dependent variable for two

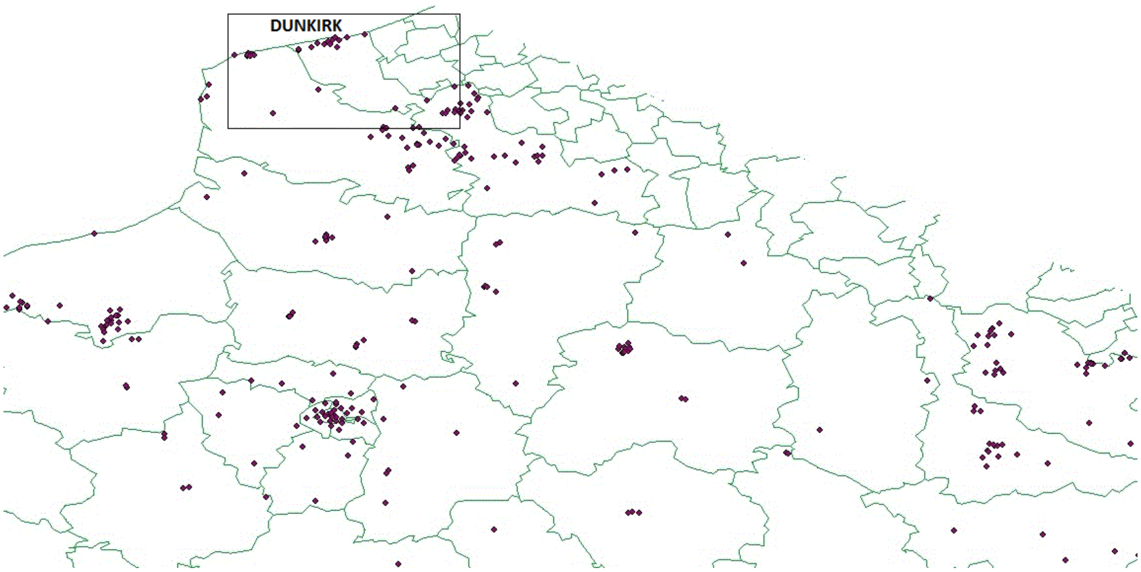
2. There is still a difference in observations because I face some missing data.

Table 4.1: Summary statistics Mean[SE]

Variables	The entire period	Before sept. 2009	After sept. 2009
	Panel A : Pollution concentration N=185,687		
Sulfur dioxyde(SO_2) ($\mu g/m^3$) mean average	2.120748[1.787319]	2.478793 [1.995779]	1.485385[1.078837]
	Panel B : Health outcomes		
Number of respiratory admission by inseecode, month, age and sex	.0586082 [.2804725]	.0568436 [.276021]	.060765 [.2858043]
Length of stay(days)	6.140569 [8.302577]	6.173942 [8.536914]	6.102398 [8.026011]
	Panel C : Weather variables		
Precipitations (mm)	2.008765 [1.015537]	2.087354 [.8599679]	1.869306 [1.232342]
Max_Temp (^0C)	14.59093 [6.513577]	14.25712 [6.044898]	15.18327 [7.233453]
Av_Temp (^0C)	10.73392 [5.627273]	10.4839 [5.27849]	11.17759 [6.173196]
wind_speed (m/sec)	7.827492 [1.123382]	7.956915 [1.17471]	7.597826 [.9849877]
Wind_direct (rose des vents)	206.6803 [44.91593]	211.3447 [40.79121]	198.4032 [50.37515]
Min_Humidity (%)	60.00654 [11.78418]	60.55584 [10.71448]	59.0318 [13.41971]
Max_Humidity(%)	93.76189 [2.770849]	92.83191 [2.65916]	95.41219 [2.120282]
	Panel D : socio-economic variables		
Age (in days)	13389.47 [12241.85]	13389.47 [12241.86]	13389.47 [12241.86]
Unemployment (%)	12.2314 [2.432976]	11.5939 [2.346794]	13.01057 [2.3063]
	The entire period	Before 2010	After 2010
	Panel E : Flat variables N=2848		
price_ttc (Euros)	128914.1 [63255.87]	126400.6 [59737.58]	131217.9 [66253.21]
level_number (number)	1.958224 [1.745731]	1.800306 [1.711344]	2.111111 [1.765608]
typ_flat (dummy)	.7847612 [.4110601]	.8193833 [.3848416]	.7530283 [.4313953]
terrace(dummy)	.1264045 [.332363]	.1666667 [.3728149]	.089502 [.2855631]
attic (dummy)	.0551264 [.2282669]	.0594714 [.2365917]	.051144 [.2203656]
balcon (dummy)	.2247191 [.4174705]	.2129222 [.4095234]	.2355316 [.4244735]
parking (dummy)	.6646106 [3.454972]	.9608856 [4.093888]	.3405973 [2.543684]
garden(dummy)	.0400281 [.1960595]	.041116 [.1986314]	.039031 [.1937338]
house_srf (m^2)	61.05585 [27.84797]	61.59224 [26.06801]	60.57463 [29.35488]
less_5_years (dummy)	.2977528 [.4573505]	.3230543 [.467815]	.2745626 [.4464438]
room_nb (number)	2.594101 [1.308262]	2.709251 [1.20337]	2.48856 [1.389549]
	Panel F : House variables N=13870		
price_ttc (Euros)	152684.1 [97988.77]	150242.5 [68782.41]	154983.5 [119076.9]
pool (dummy)	.0021629 [.0464588]	.0017839 [.0422012]	.0025199 [.0501393]
typ_house (dummy)	.3647441 [.4813757]	.4148952 [.4927406]	.3175136[.4655417]
terrace (dummy)	.0487383 [.2153281]	.0404341[.1969897]	.0565589 [.2310139]
attic(dummy)	.1312906 [.33773]	.132154 [.3386833]	.1304774 [.3368515]
balcon(dummy)	.0023071 [.047979]	.0019325[.0439211]	.0026599 [.0515096]
parking(dummy)	.2772987 [3.591735]	.2660619[3.460214]	.2879751 [3.712596]
house_srf (m^2)	107.198 [42.3946]	106.4439[40.30485]	107.9031[44.25248]
less_5_years (dummy)	.0515501 [.2211249]	.0567861 [.2314506]	.0466191 [.2108363]
room_nb (number)	3.888745 [2.429112]	3.998365 [2.3414]	3.785494[2.504704]

Note: This table indicates the mean and standard error for the estimation key variables from 2008 to 2011 in France, before shutting down the refining process and after the shutting down in the north of France.

Figure 4.4: Monitoring stations and the Nord-Pas de Calais region



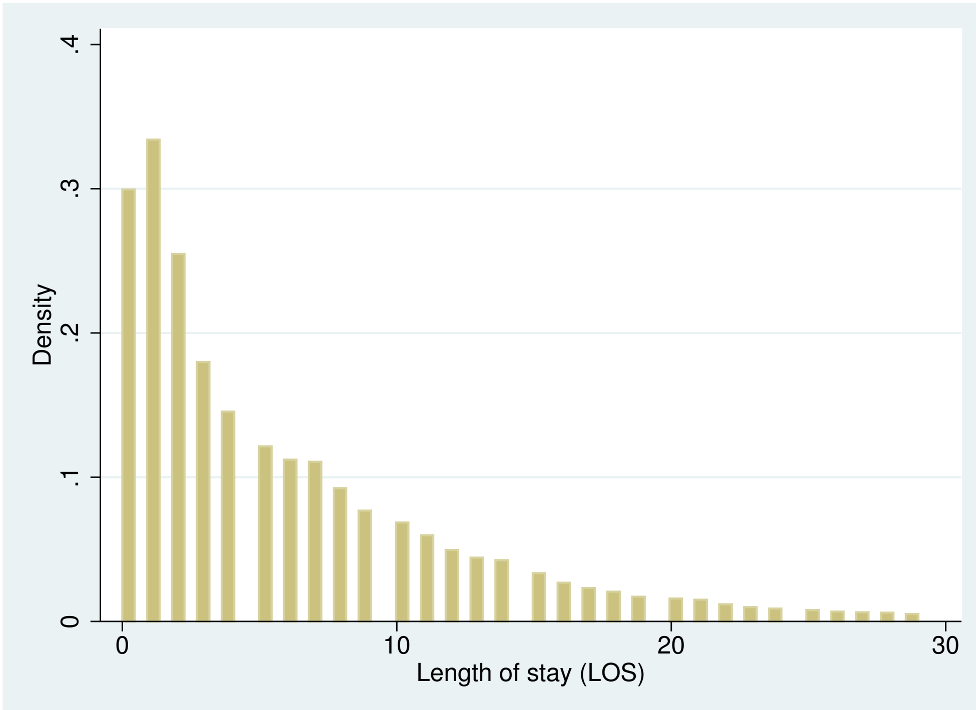
Note: This figure represents the distribution of monitoring stations in France. The marquee sheds light on the area of the study.

independent groups. The results indicate that there is a statistically significant difference between SO_2 concentration for the control group and Dunkirk ($p = .000$). In other words, Dunkirk has a statistically significantly higher SO_2 concentration (9.920242) than the control group (2.725086).

4.3.2. Morbidity data

Health data are drawn from the French National Hospital Discharge Database (PMSI) from 2008 to 2011. The key variables for the analysis are the month of admission, the length of stay and the place of residence of the patient. I dispose of an anonymous summary which gives information about the geographical code of residence of a patient, its age, its sex, its main and linked diagnosis. Pathologies are classified with respect to the international disease classification. I dispose of both outpatient discharges and emergency discharges. People who did not stay overnight in the hospital have a length of stay of zero in the dataset. I do not dispose of the exact day of admission but I have the length of stay (LOS). I relied on this information to construct a measure of expected exposure to air pollution. I only dispose of the month of discharge (not month of admission), so I cannot precisely determine exposure to the closure. To address this, I combine information on the month of discharge with LOS to assign probabilities of admission to weight observations. More specifically, based on a patient's LOS, I assign two possible months of admission, doubling the number of observations. I then perform a weighted regression where I weight each observation by the probability of being admitted in that month as explained in more details in Annex 1. For example, for a patient discharged in October with an LOS of 10, there are only two possible months of admission: September and October. The probability of admission in September is 10/30 and in October is 20/30 (=1-10/30). I then weight each observation by these probabilities, and perform the regressions of these weighted values. To gauge how important this may be, Figure 4.5 shows the distribution of LOS. As expected in this plot, nearly all patients (90%) are in the hospital for less than 15 days. Therefore, despite not knowing the month of admission, I suspect this will not significantly bias our results. To assess the degree to which this may occur, I estimate models that assume the month of admission is the month of discharge, and compare obtained results to that of the probability weighted approach.

Figure 4.5: The distribution of the length of stay (LOS)



Note: This graph represents the density of probability for the length of stay for stay below 30 days. The length of stay is in number of days.

Panel B of the summary statistics sheds light on the number of admissions for respiratory disease by month, year, municipality, age in days and sex to keep the most disaggregated dataset. Thus, the unit of observation is an imputed person-month of exposure. Note that most of the dataset consists of patients that have stayed less than 15 days (90 percent) at the hospital. Length of stay (LOS) after the refinery closure in the treated group has decreased from 7.09 to 6.06 days whereas LOS in the other municipalities, has decreased from 6.04 to 5.99 days in average. As a consequence, the difference in difference between Dunkirk and the control group after the refinery closure is about 1 day.

4.3.3. Weather data and socioeconomic data

I use temperature, precipitation, humidity and wind data in the analysis to control for both the direct effects of weather on health (Chay & Greenstone, 2003) and also to leverage the quasi-experimental features of wind direction and wind speed in distributing pollution from refineries (Hanna & Oliva, 2011), (Beatty & Shimshack, 2011). The weather data comes from Meteo France, the French national meteorological service. I dispose of the average and maximum temperature in Celsius degree, the number of precipitation in millimeters, the maximum speed wind in meters per second, the prevailing wind direction in wind rose and the maximum and minimum relative humidity in percent.[3] I use daily data from the French weather monitoring system. I use measures from one nearby station per department as a control in the regression. Weather data are presented in Panel C of the summary statistics.

Temperature and the intensity of sunlight play an important influence in the chemical reactions that occur in the atmosphere to form photochemical smog from other pollutants. Besides, wind speed and direction measurements are important for air quality monitoring. If high pollutant concentrations are measured at a monitoring station, the wind data recorded at the station can be used to determine the general direction and area of the emissions. Wind speed can greatly affect the pollutant concentration in a local area. The higher the wind speed, the lower the pollutant

3. The relative humidity of an air-water mixture is defined as the ratio of the partial pressure of water vapor in the mixture to the saturated vapor pressure of water at a prescribed temperature.

concentration. Wind dilutes pollutants and rapidly disperses them throughout the immediate area. Humidity and precipitation can also act on pollutants in the air to create more dangerous secondary pollutants, such as the substances responsible for acid rain. On the opposite, precipitations have a beneficial effect by washing pollutant particles from the air and helping to minimize particulate matter formed by activities such as construction and some industrial processes (Environmental Protection Agency (EPA), 2012a).

I also use the quarterly rate of unemployment from the National Institute of Statistics and Economic Studies (INSEE). 30% of flats in the region are less than 5 years old. In December 2008, a new law, known as the "Scellier" Act, was passed (a measure adopted under the 2008 corrective finance Act 2008-1443 of December 30, 2008), which gives taxpayers who acquire a new home or a home before completion between January 1, 2009 and December 31, 2012, subject to various conditions, a reduction on their income tax of 25% to 37% of the cost of the purchase, spread over a period of 9 to 15 years. The new incentive resulting from the Scellier Law is intended to give a boost to rental properties and increase the attractiveness of an investment in new construction.

4.3.4. Property Prices

I use a unique and really rich dataset coming from the Chambre des notaires, PERVAL dataset. I dispose of every property transactions for a girth of 50 km around Dunkirk from 2008 to 2011. The geographic region for the control group is then represented by all towns with their centers within 50 km of the refinery in Dunkirk. The control variables that are used in the estimation have really few missing values in order to keep the most exhaustive dataset. Panel E and F of Table 4.1 present the main characteristics of the 2848 and 13870 property transactions for flats and houses respectively. The key variables are the property prices, the number of floors, the number of rooms, the type of flat or house, the property surface, the presence of a terrace, an attic, a parking, a balcony, a pool or a garden and a variable which indicates if the property is less than 5 years old. I observe from the summary statistics that property prices for both houses and flats increase on average after the refinery closure.

4.4. Estimation

I am looking at the causal relationship between a closure in the refining activity, local pollution levels, the number of contemporaneous respiratory hospitalizations, and the property prices in the North of France.

The purpose, first, is to estimate the impact of the refining closure $post\_closure_{cm}$ on pollution concentration or health outcomes Y_{cm} captured by the parameter β_1 in the following linear probability model:

$$Y_{cm} = \beta_0 + \beta_1 post\_closure_{cm} + \beta_2 post_{cm} + \beta_3 closure_{cm} + X_{cm} + \theta_m + \omega_y + \epsilon_{cm} \quad (4.1)$$

where the dependent variable Y_{cm} represents either SO_2 pollution concentration or health outcomes within each municipality c at month m according to the model I present. The variable $post_{cm}$ represents the timing after the refinery closure in september 2009. I also include in all regressions the variable $closure_{cm}$ coded as 1 for Dunkirk to control for time-invariant unobserved covariates of respiratory admissions. The control group coded as zero represents the rest of municipalities in the north of France with at least one monitoring station. The variable $post\_closure_{cm}$ is the difference in difference estimator and represents Dunkirk area after the refinery closure in 2009. Because I only dispose of monthly hospital admissions, I aggregate the daily measure of pollution concentration at a monthly level.[4] X_{cm} is a vector of municipality controls that includes weather controls W_{cm}. I also control for temporal variation in pollution including month fixed effects θ_m, year fixed effects ω_y to limit the influence of pollution outliers. Finally, ϵ_{cm} represents the error term.

This difference in difference model assumes the following for the model to be valid:

$$\mathbb{E}[\epsilon_{cm}, post\_closure_{cm}] = 0 \tag{4.2}$$

The refinery in Dunkirk ceased its activity in September 2009, and in October 2010, the owner, Total, announced the facility would be permanently closed. From a pollution standpoint, emissions ceased in September 2009, that is why we first use a monthly model when estimating the impact of pollution concentration on health. The chemical life of SO_2 is around 2 days, well below the period of time of

4. When aggregating daily data, the monthly average concentration take into account monitors without missing data

the study (a month) which removes any problems of persistence or autocorrelation. If it was the length of exposure which was responsible of a pathology, I could risk identifying a change in the population composition instead of the effect of the refinery closure. In fact, employees from the refinery, very exposed to pollution could leave Dunkirk following the activity stop leading to a reduction in the number of hospital admissions without any effect on health. However, the impact of SO_2 is relatively local with a direct effect on health such that the concentration from month m does not depend on concentration from the previous month m-1. The exogenous reduction in SO_2 in Dunkirk is responsible for the change we can observe in the number or the severity of an admission. The effect of the refinery closure is well identified.

Similarly, the second objective is to estimate the impact of the refining closure $post\_closure_{pcy}$ on housing prices P_{pcy} or the parameter α_1 in the following linear probability model:

$$P_{pcy} = \alpha_0 + \alpha_1 post\_closure_{pcy} + \alpha_2 post_{pcy} + \alpha_3 closure_{pcy} + W_{pcy} + \varphi_y + \nu_c + \sigma_{cy} \quad (4.3)$$

where the dependent variable P_{pcy} represents the log price of each property p, within each municipality c at year y. The variable $post\_closure_{pcy}$ represents Dunkirk after the refinery closure in 2009. A semi-log functional form is the most preferred because it fits the data particularly well (Palmquist, 1984). W_{pcy} is a vector of property and municipalities controls that include property characteristics and the level of unemployment in each area. I also control for temporal variation in pollution including year fixed effects η_y to limit the influence of pollution outliers. I also include in all regressions a municipality fixed effects ν_c to control for time-invariant unobserved covariates of prices. In fact, much of the variations may be explained by unobserved factors that characterize particular properties like geographical features, neighborhood characteristics and design amenities (Davis, 2004). σ_{cy} represents the error term.

One of the fundamental assumptions of the hedonic price method, as underlined previously, is that households have perfect information. If people are not fully informed of the linkages between the environmental attributes and benefits to them or their property, the value will not be reflected in home prices. Moreover, the hedonic price schedule does not adjust instantaneously to changes in demand or supply conditions in the housing market. Many factors like imperfect information and transaction costs will then result in the process of adjustment taking some time. In addition, the possibility that the refinery might be definitely closed was evident for some time, due to several justices decision from the beginning of 2010. Before October 2010 people might have indeed suspected that the temporary shut-down would be permanent after September 2009. That is why I assume a buying decision being made over a year in this model such that I observe the change in housing prices from 2010, three months after the refinery activity stop and ten months before the real closure. In other words, the model compares 2010 and 2011 versus 2008 and 2009. I reasonably assume that 2010 takes into account the delay for adjustment to changes in demand and the anticipation of the closure people can have before the refinery definitely shutted down.

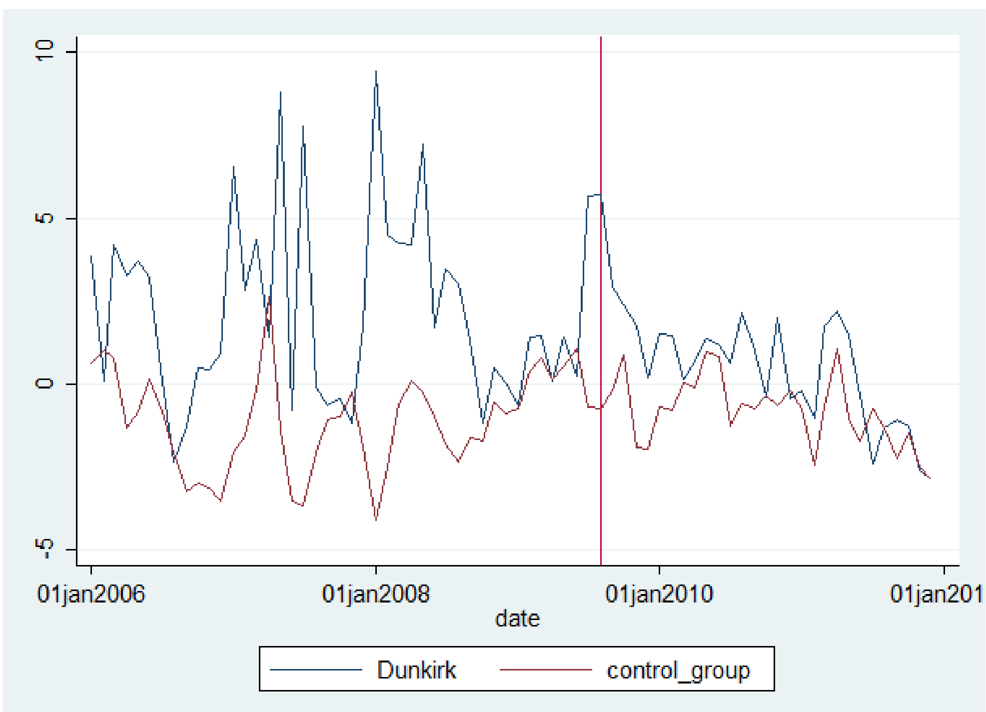
Besides, the issue of sorting is crucial in this study. Individuals that choose to live near toxic sites may have a low willingness to pay to avoid the associated health risks. If consumers value the closure of the refinery, then the closure should cause individuals to sort themselves such that there is an increase in the number of people, who place a high value on environmental quality, living near the refinery. Hanna & Oliva (2011)'s paper supports that the closure may have altered the attractiveness of surrounding neighborhoods for individuals with strong preferences for air quality. If wealthier or healthier people moved closer to the refinery after it closed, the estimates might simply be capturing the differences in terms of population characteristics between the old and new residents of the refinery neighborhood. Although focusing on the years around the closure reduces the probability of selective migration, I will analyze the change in population after the closure.

4.5. Results

4.5.1. Pollution concentration and refinery closure

I start by examining the effect of refinery closure on air pollution. The result is interesting in itself to understand to what extent refining activity influences the amount of pollution released in the air. Figure 4.5 provides a monthly graph of SO_2 residual estimation in $\mu g/m^3$ from 2008 to 2011 for Dunkirk and municipalities far from Dunkirk. I restrict the sample to the Nord-Pas de Calais region. I control for seasonal patterns adding year and month dummies to deal with the falling pollution trend and the high variation in air pollution I observe overtime. After the closure of the refinery in Dunkirk, pollution in SO_2 falls in Dunkirk relative to their counterparts where pollution seems to be stable. After the closure, the level of pollution concentration in Dunkirk catches up the level of pollution observed in the other municipalities.

Figure 4.6: Monthly residual regression for SO_2 in ($\mu g/m^3$) from 2006 to 2011



Note: This graph represents the SO_2 residual concentration for municipalities with a refinery versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure

Table 4.2 details this effect more carefully. I present the estimate of β_1 from

Equation (1), where I replace the *post_closure* variable by a dummy variable for whether or not the municipality is Dunkirk after september 2009. I consider a simple measure of SO_2 in Micrograms per cubic meter in column 1. While column 2 adds weather control, column 3 repeats the estimation with municipalities fixed effects. Column 4 adds the unemployment variable as a proxy for the economic activity trend. In the last column, I take full advantage of the variation in distance between the municipality and the monitoring station by reducing the dataset to two kilometers distance between the centroid of the municipality and monitors in order to have a more precise measure of populations' exposure to air pollution. Note that all specifications include month, year and municipality fixed effects and are clustered by municipality and month.

Dunkirk, the municipality where the refinery is located, shows a reduction in SO_2 air pollution after the refinery closure of 5 micrograms per cubic meter on average. The refinery closure substantially reduces pollution and it is consistent with all measures of pollution. The estimate is not driven by standard demographic characteristics (column 3) nor by neighborhood specific trends (municipality FE). Taking distance into account increases the magnitude of the effect and the estimate remains significant at the five percent level.

4.5.2. Respiratory outcomes and refinery closure

I now focus on the health impact evaluation of the refinery closure. Results from the reduced form are important because refinery closure may involve benefits

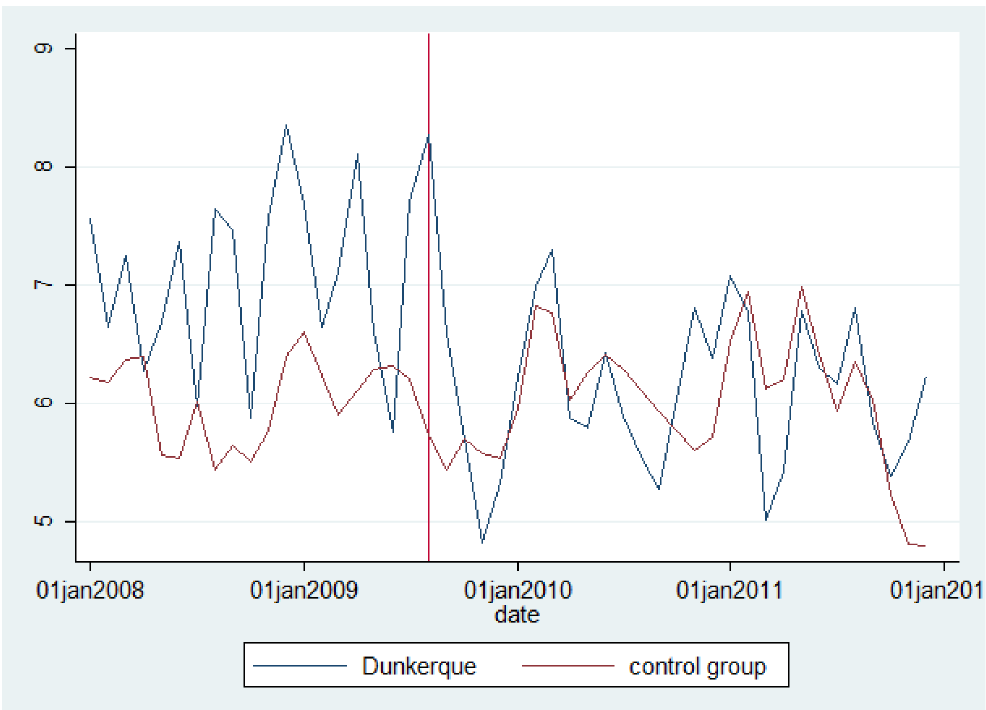
Table 4.2: First stage regression

VARIABLES	(1) SO_2	(2) SO_2	(3) SO_2	(4) SO_2	(5) SO_2
post_closure	-5.166**	-5.163**	-4.951**	-4.993**	-5.072**
	(2.146)	(2.150)	(2.166)	(2.188)	(2.214)
post	0.856	1.059			
	(0.627)	(0.643)			
closure	9.547***	9.536***			
	(2.065)	(2.060)			
av_temp		-0.764	-0.358	-0.316	-0.221
		(0.431)	(0.439)	(0.447)	(0.462)
pp		-0.192	-0.179	-0.199	-0.243
		(0.156)	(0.156)	(0.163)	(0.164)
max_temp		0.734	0.365	0.348	0.283
		(0.417)	(0.415)	(0.420)	(0.432)
speed_wind		0.362**	0.321*	0.339**	0.405**
		(0.129)	(0.149)	(0.152)	(0.158)
direct_wind		-0.00301	-0.00306	-0.00324	-0.00395
		(0.00473)	(0.00518)	(0.00521)	(0.00540)
min_humidity		0.00528	-0.0242	-0.0209	-0.0252
		(0.0411)	(0.0414)	(0.0409)	(0.0383)
max_humidity		0.0214	0.0306	0.0461	0.0654
		(0.0886)	(0.0948)	(0.102)	(0.101)
Un				-0.327	-0.317
				(0.251)	(0.253)
Year FE	x	x	x	x	x
Month FE	x	x	x	x	x
municipalities FE			x	x	x
Distance < 2km					x
Observations	185,687	185,687	185,687	185,687	175,212
Adjusted R-squared	0.478	0.486	0.503	0.504	0.504

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the concentration of SO_2. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

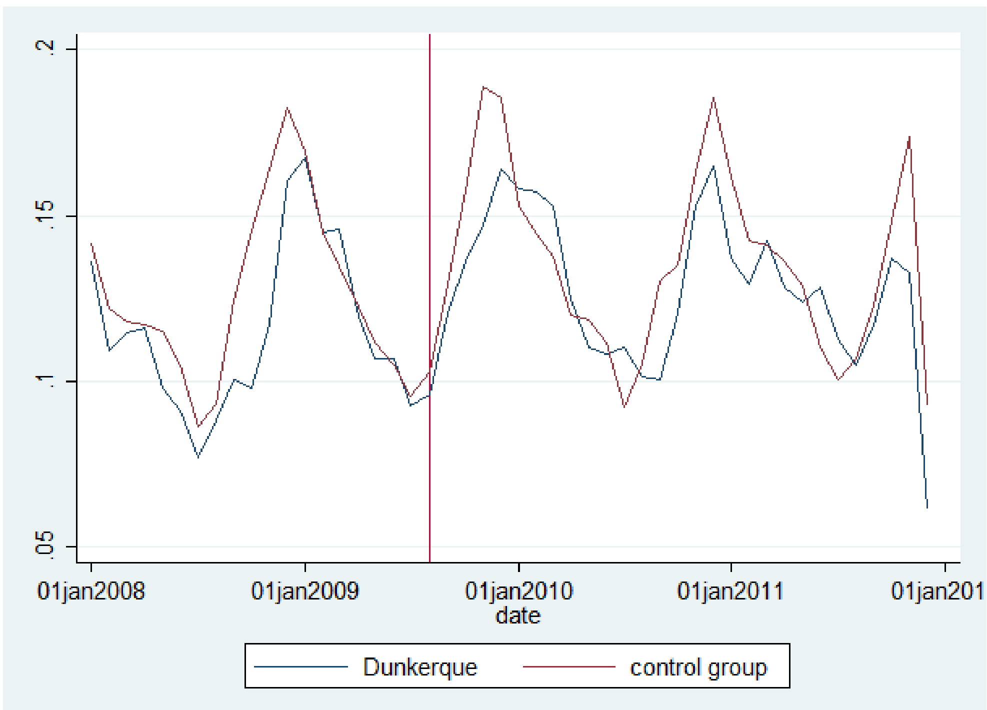
(Hanna & Oliva, 2011) and may help to establish the extent to which refinery closure reduces economic costs in the form of loss of earning profits and wages. Figure 4.7 and Figure 4.8 provide monthly graphs of the evolution of the number of respiratory admission and the evolution of the length of stay at the hospital during the period respectively.

Figure 4.7: Evolution of the lengh of stay in days from 2008 to 2011



Note: This graph represents the evolution of the length of stay for the area of Dunkirk versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure.

Figure 4.8: Evolution of the number of admissions from 2008 to 2011



Note: This graph represents the evolution of the number of admissions for the area of Dunkirk versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure.

By disentangling the extensive from the intensive margin, the study shows to what extent SO_2 air pollution triggers a disease or increases its severity. It is quite clear from the graph that the length of stay in Dunkirk falls after the closure of the refinery in such a way that it closes the gap between Dunkirk and the rest of the municipalities. Concerning the number of admissions, I do not observe any obvious changes from the graph. One should note the number of respiratory admissions is higher in winter due to the relationship between temperature and respiratory admissions. I present the estimate of β_1 from Equation (4.1) in table 4.3, 4.4 and 4.5. Note that table 4.3 uses a logit model. The outcome is coded 1 if there is at least one hospital respiratory admission and zero otherwise.

Adults could adjust their behavior after the closure of the refinery. If they remain unemployed after the refinery closure, they may have more time to go to the hospital for a visit. The length of stay may be a better outcome to control for any avoidance behavior as they have no influence on the number of days they will stay at the hospital once they are admitted. Tables 4.4, 4.5, 4.6 indicate results for the length of stay, the unit of analysis being an admission.

As I cannot rule out the possibility of avoidance behavior by workers because of reductions in hours worked, I do not denote any decline in the number of admissions for respiratory outcomes in table 4.3. Introducing the unemployment variable sharply reduces the coefficient on the post variable, presumably because unemployment trends upward over the period within the entire region. I find a slowdown in

Table 4.3: Respiratory admissions: reduced form regressions

VARIABLES	(1) Admissions	(2) Admissions	(3) Admissions	(4) Admissions
post_closure	0.0216	0.0219	-0.00822	-0.0113
	(0.0396)	(0.0395)	(0.0389)	(0.0398)
post	0.0989***	0.0767***	0.182***	0.184***
	(0.0198)	(0.0277)	(0.0390)	(0.0396)
closure	0.00615	0.00571	-0.0289	-0.0227
	(0.0302)	(0.0301)	(0.0311)	(0.0312)
Un			-0.0883***	-0.0930***
			(0.00503)	(0.00453)
av_temp		0.00362	-0.0166	-0.00798
		(0.0478)	(0.0433)	(0.0412)
pp		-0.00505	-0.00790	-0.00523
		(0.0152)	(0.0167)	(0.0180)
max_temp		0.00662	0.0274	0.0199
		(0.0423)	(0.0402)	(0.0386)
speed_wind		0.00887	0.0139	0.00931
		(0.0134)	(0.0164)	(0.0187)
direct_wind		-1.92e-05	-0.000136	-0.000116
		(0.000314)	(0.000305)	(0.000320)
min_humidity		0.00243	0.00548	0.00453
		(0.00495)	(0.00566)	(0.00592)
max_humidity		-0.00279	0.00179	0.00427
		(0.00885)	(0.0114)	(0.0115)
Year FE	x	x	x	x
Month FE	x	x	x	x
Distance < 2km				x
Observations	185,687	185,687	185,687	175,212

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on respiratory admissions. All regressions are estimated using logit, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

the severity of illness in table 4.4 suggesting avoidance behavior is unlikely to drive our results. Length of stay (LOS) is reduced in average by one day after the refinery closure.

Table 4.4: The length of stay: reduced form regressions

VARIABLES	(1) LOS	(2) LOS	(3) LOS	(4) LOS	(5) LOS
post_closure	-1.083***	-1.082***	-1.087***	-1.096***	-1.104***
	(0.257)	(0.267)	(0.244)	(0.246)	(0.234)
post	-0.138	-0.448			
	(0.336)	(0.355)			
closure	1.062***	1.055***			
	(0.209)	(0.221)			
av_temp		0.269	0.169	0.192	0.184
		(0.327)	(0.328)	(0.326)	(0.362)
pp		-0.255**	-0.240**	-0.249**	-0.281**
		(0.111)	(0.0986)	(0.0981)	(0.111)
max_temp		-0.267	-0.180	-0.191	-0.198
		(0.306)	(0.313)	(0.308)	(0.337)
speed_wind		0.213	0.188	0.194	0.253*
		(0.138)	(0.139)	(0.138)	(0.127)
direct_wind		-0.00116	-0.000673	-0.000725	-0.00184
		(0.00239)	(0.00260)	(0.00261)	(0.00277)
min_humidity		-0.0187	-0.0181	-0.0167	-0.0127
		(0.0364)	(0.0397)	(0.0393)	(0.0420)
max_humidity		0.0127	0.0195	0.0261	0.00988
		(0.0421)	(0.0442)	(0.0451)	(0.0496)
Un				-0.146	-0.139
				(0.200)	(0.202)
Year FE	x	x	x	x	x
Month FE	x	x	x	x	x
municipalities FE			x	x	x
Distance < 2km					x
Observations	18,544	18,544	18,544	18,544	17,514
R^2	0.003	0.003	0.005	0.005	0.006

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the number of respiratory admissions. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Table 4.5 repeats the same procedure, doubling the number of observations and using the length of stay as a weight to take into account the uncertainty about the month of admission. The decrease in concentration of air pollution may not be sufficient to reduce the number of respiratory diseases whereas it may be enough to have an effect on the severity of a respiratory disease.[5]

As previously indicated, children and the elderly are particularly sensitive to pollution. To assess this, I repeat the same regressions stratifying by age group in table 4.6 taking again into account the uncertainty about the month of admission. This assessment also helps to probe into potential avoidance behavior.

5. The estimation of the length of stay may present a selection bias. In fact, it only considers people already admitted at the hospital. However, the non significativity I observe in the estimation of the admission shows that there is no extensive margin effect

Table 4.5: The length of stay: reduced form regressions with respect to the probability of admission

VARIABLES	(1) LOS	(2) LOS	(3) LOS	(4) LOS	(5) LOS
post_closure	-1.16***	-1.16***	-1.15***	-1.16***	-1.16***
	(0.190)	(0.201)	(0.194)	(0.194)	(0.184)
post	-0.02	-0.23			
	(0.234)	(0.279)			
closure	1.07***	1.06***			
	(0.171)	(0.176)			
av_temp		0.41	0.38	0.40	0.43
		(0.279)	(0.266)	(0.258)	(0.259)
Pp		-0.25	-0.25*	-0.26*	-0.28*
		(0.145)	(0.134)	(0.136)	(0.141)
max_temp		-0.43	-0.40	-0.41	-0.45*
		(0.265)	(0.251)	(0.245)	(0.246)
speed_wind		0.12	0.10	0.11	0.16
		(0.118)	(0.108)	(0.110)	(0.105)
direct_wind		-0.00	-0.00	-0.00	-0.00
		(0.00201)	(0.00204)	(0.00204)	(0.00195)
min_humidity		-0.02	-0.02	-0.02	-0.03
		(0.0265)	(0.0255)	(0.0252)	(0.0251)
Max_humidity		0.05	0.05*	0.06*	0.06*
		(0.0323)	(0.0288)	(0.0303)	(0.0318)
Un				-0.14	-0.14
				(0.146)	(0.149)
Year FE	x	x	x	x	x
Month FE	x	x	x	x	x
municipalities FE			x	x	x
Distance < 2km					x
Observations	34,641	34,641	34,641	34,641	32,690
R^2	0.002	0.003	0.005	0.005	0.005

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the number of respiratory admissions taking into account the uncertainty about the month of admission. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Table 4.6: LOS Reduced form regressions by age category with respect to the probability of admission

Variables	(1) 0-5	(2) 5-15	(3) 15-25	(4) 25-40	(5) 40-60	(6) 60-70	(7) >70
post_closure	-0.260*	-0.203	0.536	-1.631*	-0.986	-0.851	-2.127***
	(0.122)	(0.578)	(0.683)	(0.744)	(0.759)	(1.048)	(0.490)
Un	0.160	0.0389	-0.159	-0.273	0.389	-1.755***	0.222
	(0.0904)	(0.196)	(0.529)	(0.290)	(0.227)	(0.478)	(0.269)
av_temp	0.148	0.218	-0.104	1.206**	0.418	1.477*	0.515
	(0.177)	(0.321)	(0.311)	(0.432)	(0.872)	(0.738)	(0.510)
pp	-0.119	-0.0678	-0.0497	-0.277	-0.185	-0.572***	-0.464*
	(0.0701)	(0.0826)	(0.181)	(0.200)	(0.235)	(0.175)	(0.220)
max_temp	-0.164	-0.203	0.0139	-1.104**	-0.325	-1.461*	-0.519
	(0.165)	(0.291)	(0.325)	(0.471)	(0.774)	(0.686)	(0.493)
speed_wind	0.0529	-0.0388	0.110	-0.0885	0.0313	0.571**	0.0938
	(0.0872)	(0.109)	(0.176)	(0.274)	(0.251)	(0.236)	(0.175)
direct_wind	0.00153	-0.00174	0.00151	0.00302	-0.00162	-0.00699	-0.00239
	(0.00220)	(0.00312)	(0.00354)	(0.00452)	(0.00321)	(0.00633)	(0.00363)
min_humidity	-0.0148	-0.0189	0.0375	-0.0704	0.0260	-0.121*	-0.0408
	(0.0186)	(0.0481)	(0.0520)	(0.0818)	(0.0619)	(0.0593)	(0.0591)
max_humidity	-0.00741	0.0535	-0.155	0.129	-0.00800	0.146	0.0700
	(0.0312)	(0.0944)	(0.0887)	(0.146)	(0.0835)	(0.102)	(0.0916)
Year FE	x	x	x	x	x	x	x
Month FE	x	x	x	x	x	x	x
Municipalities FE	x	x	x	x	x	x	x
Observations	9,205	1,958	1,882	3,035	6,023	3,375	9,163
R^2	0.024	0.045	0.025	0.025	0.008	0.035	0.016

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the number of respiratory admissions by age category. All regressions are estimated using OLS, with standard errors clustered at the month and department level. Demographic controls include an indicator for gender, age and unemployment. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

Consistent with this, I only find statistically significant estimates for the youngest (under age 5) and oldest (>70) age group. These differences by age also help to allay concerns regarding avoidance behavior. While workers may have dramatically changed their exposure because they were no longer working, it is unlikely that children changed their behavior. And since nearly the entire elderly group is not working, it is unlikely their behavior changed either. Hence, these results suggest that for at least some groups I am picking up effects net of avoidance behavior. Thus, the estimates suggest that daily variation in SO_2 air pollution has economically significant effects on the severity of respiratory outcomes for the at risk population. The stop in activity also reduces the length of stay at the hospital for people between 25 and 40 years old. Perhaps surprisingly, I find a statistically significant decrease in the LOS for people ages 25-40. This last category is part of the working population and could represent refinery workers. The refinery closure represents an economic shock. Chay and Greenstone (2003) state a decline in business activity may affect pollution level and employment patterns. In this context, we could think that the health of refinery workers has been directly impacted by an air pollution improvement. Moreover, I am in line with the literature saying that pollution significantly reduces worker productivity through health. Thus, cleaning up pollution from refining activity would benefit the economy in terms of labor. This result suggests that workers with young children are even more likely to miss a day of work because pollution impacts their own health and their children's. However, because air pollution levels are geographically determined, it is very difficult to assess exposure-response relations at an individual level. Exposure differences between age categories may induce measurement error. For example, the effect may be stronger for 25-40 years old population compared to 40-60 years old population due to differences in the time spent outdoors.

4.5.3. Property prices and refinery closure

I have previously underlined a causal effect between a reduction in pollution concentration and a reduction in the severity of illness. I now wonder how people perceive this reduction in SO_2 concentration and the improvement in health status. This reduction in pollution and health risk after the refinery closure may also have some effect on property values, reflecting perception and WTP, that I investigate in this section. An overall property market model is first presented. Then, the dataset is divided in two separated segments due to their specificity and differences in price evolution: flats and houses. Buying a flat is not the same investment decision as buying a house, and the reasons behind such an investment may differ.

Overall property market In these following property models, the exclusion restriction may be violated if other events happen at the same time of the closure. Nevertheless, it is unlikely that this is the case. As far as I know, no other economic events or policies coincided with that particular period. However the area of Dunkirk is exposed to many economic activity downturns. The economic activity evolves constantly in Dunkirk mainly due to its port. France's third-ranking port, Dunkirk is well known as a port handling heavy bulk cargoes for its numerous industrial installations. Dunkirk represents also the France's leading port for ore and coal imports; the France's leading port for containerized fruit imports; France's leading port for copper imports; and the France's second-ranking port for trade with Great Britain. The port's territory covers 7,000 hectares and includes ten towns : Dunkirk, Saint-Pol-sur-Mer, Fort-Mardyck, Grande-Synthe, Mardyck, Loon-Plage, Gravelines, Craywick, Saint-Georges-sur-l'Aa and Bourbourg. Thus, I control in every estimation for unemployment, as a control variable for the economic activity. Table 4.7 shows that the refinery closure does not have any impact on unemployment at a census tract level, because the coefficient is not significant. After the closure, jobs have indeed been offered to the workers of the refinery in other group facilities or unit. Table 4.7 also introduce *buyer_migration*, *single* and *male* outcome variables which are coded as one if buyers migrate outside Dunkirk, are single and male respectively. The refinery closure does not have any impact on the composition of the population as I do not find any significance for those coefficients.

Table 4.7: Other effect for the overall sample

VARIABLES	(1) Migration_buyer	(2) Un	(3) Male	(4) Single
post_closure	-0.0197	0.0317	-0.00468	-0.000505
	(0.0277)	(0.125)	(0.0182)	(0.0315)
Year FE	x	x	x	x
municipality FE	x	x	x	x
Observations	16,718	16,159	16,718	16,718
Adjusted R-squared	0.113	0.964	0.169	0.054

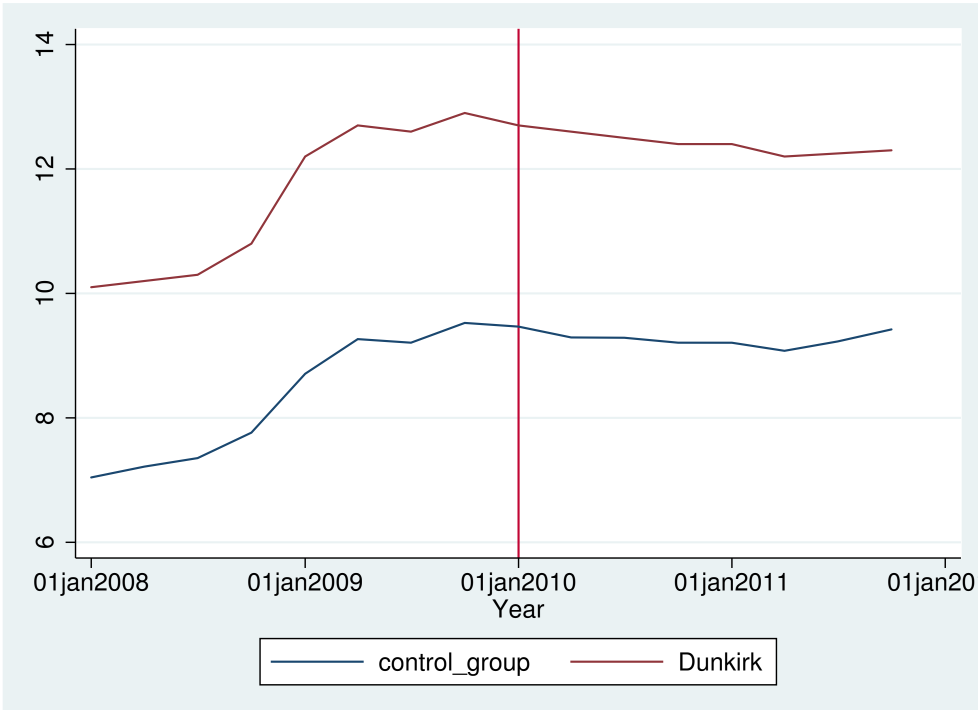
Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on demographics. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

To reinforce the assumption of homogeneity between the treated and the control group, it is necessary to stress that one municipality in the control group includes Calais port. Located on the busiest straits in the world for international shipping, the port of Calais alone handles 1/3 of the traffic between continental Europe and the United Kingdom. In this context, Calais and Dunkirk represent two significant commercial ports in France. Calais as part of the control group, strengthens the existence of a similarity between the treated group and the control group. In addition, Figure 4.9 shows a similar evolution for unemployment between the control group and the treated group over the period and gives an idea about the background economic trend.[6] The unemployment time series increases over the period and serves as a time trend in the analysis. The estimation from the hedonic model suggests the proportion of the total variance that is attributed to unemployment only accounts for 0.66% in Dunkirk. It is not surprising knowing that there are only 370 workers at the refinery and those 370 workers, after the closure, were offered jobs in other group facilities or unit.

Like Currie *et al.* (2013), since in our empirical application, all residents near or far from the refinery live in the Nord-Pas de Calais region, I assume that the wage effects are similar for both nearby residents and those a little further from a plant. In this context, holding all other factors fixed, the entire change in property prices following the shock in Dunkirk compared to its counterparts indicate to what extent individuals evaluate health risk and any effect on neighborhood aesthetics.

6. Unemployment is only reported at the quarterly aggregate levels (INSEE) rather than the monthly level for accuracy reasons.

Figure 4.9: Evolution of unemployment from 2008 to 2011



Note: This graph represents the quarterly evolution of Unemployment for the area of Dunkirk versus municipalities without a refinery within 50 kilometers of Dunkirk. September 2009 corresponds to the refinery closure.

First of all, table 4.8 presents the effect of the refinery closure on the overall property market. While the first block of table 4.8 looks at the entire dataset, the second block reduces the sample to municipalities located within a distance of 10 kilometers around the refinery to have a more homogeneous dataset in terms of trends (Currie *et al.*, 2013).

Table 4.8: Reduced form regressions: overall property market

VARIABLES	(1) OLS	(2) OLS	(3) FE	(4) FE	(5) distance<10km OLS	(6) OLS	(7) FE	(8) FE
post_closure	-0.0637	-0.0143	-0.0169	-0.0199*	-0.0149	0.0409	0.0387	0.0123*
	(0.0586)	(0.0463)	(0.0125)	(0.0104)	(0.0126)	(0.0181)	(0.0190)	(0.00503)
post	0.0661	0.0279			-0.0392**	-0.0346**		
	(0.0708)	(0.0600)			(0.00673)	(0.00788)		
closure	0.0802*	0.102**			0.0238	0.0770***		
	(0.0408)	(0.0305)			(0.0101)	(0.00672)		
terrace		0.121***	0.105***	0.106***		0.0930**	0.0905**	0.0944**
		(0.0208)	(0.0213)	(0.0207)		(0.0227)	(0.0255)	(0.0256)
attic		-0.0574***	-0.0255**	-0.0285**		-0.0169	-0.0173	-0.0282
		(0.0150)	(0.00923)	(0.00874)		(0.0273)	(0.0320)	(0.0244)
balcony		0.189***	0.132***	0.133***		0.0584**	0.0616***	0.0724**
		(0.0234)	(0.0194)	(0.0184)		(0.0122)	(0.01000)	(0.0156)
parking_srf		-7.50e-05	-0.000604	-0.000812		0.00138	0.00256	0.000300
		(0.000850)	(0.00102)	(0.000968)		(0.00104)	(0.00142)	(0.000403)
garden		-0.0173	-0.0588	-0.0566		-0.0226	-0.0206	-0.00586
		(0.0393)	(0.0478)	(0.0486)		(0.0500)	(0.0522)	(0.0500)
house_srf		0.00538***	0.00527***	0.00572***		0.00485**	0.00474**	0.00677***
		(0.000384)	(0.000369)	(0.000185)		(0.00131)	(0.00114)	(0.000241)
less_5_years		0.389***	0.379***	0.382***		0.437***	0.427***	0.441***
		(0.0284)	(0.0405)	(0.0396)		(0.0244)	(0.0273)	(0.0260)
room_nb		0.0423***	0.0436***	0.0374***		0.0726**	0.0709**	0.0455***
		(0.00722)	(0.00790)	(0.00578)		(0.0160)	(0.0130)	(0.00529)
Un				-0.0294				-0.0600
				(0.0205)				(0.0320)
Year FE	x	x	x	x	x	x	x	x
municipality FE			x	x			x	x
distance<10km					x	x	x	x
Observations	16,718	10,957	10,957	10,671	4,335	2,446	2,446	2,160
Adjusted R-squared	0.003	0.401	0.490	0.502	0.003	0.468	0.504	0.562

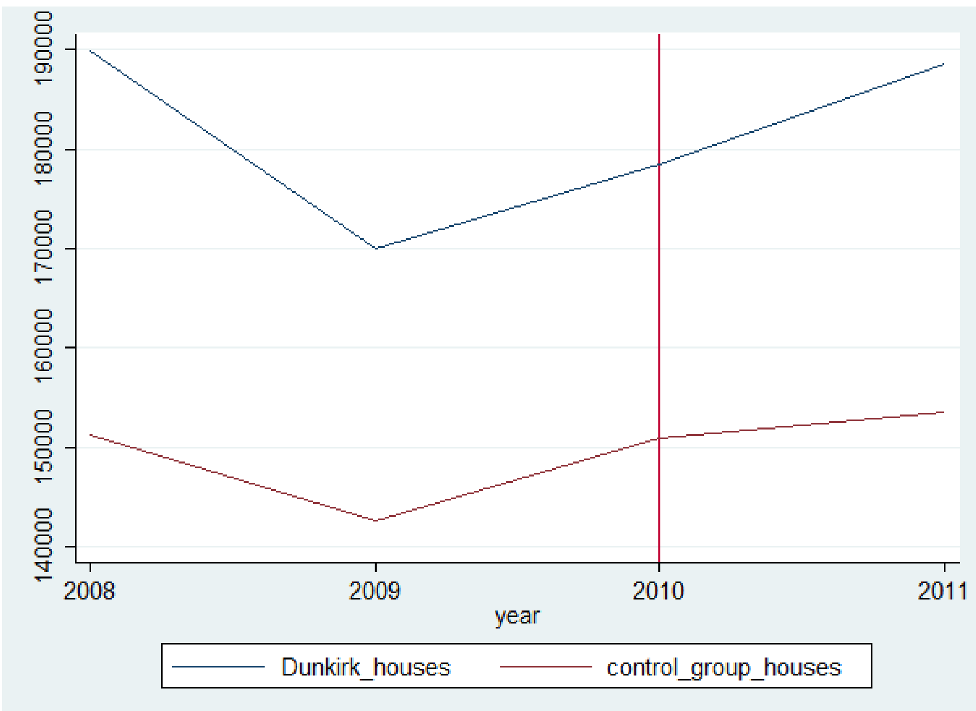
Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of house. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Column 1 uses a difference in difference model without property characteristics. Column 2 adds property characteristics and column 3 controls for municipality specificities using a municipality fixed effects. The last column takes full advantage of the unemployment variable in each municipality to control for any differences in the activity trend between the treatment and control group. Table 4.8 shows a slight decrease in the price of properties after the closure only when taking unemployment into account but it is not robust in all fixed effects models even when I present a more homogeneous dataset in the second block reducing distance. As underlined previously, flats and houses have their specificities and differences in price evolution and considering them simultaneously may strongly bias our estimates.

House Prices Figure 4.10 shows the evolution of housing prices before and after the definitive refinery closure which occurs in 2010.

The price trend in Dunkirk follows a similar increasing trend before and after the closure whereas the price trend in the others municipalities keeps rising but at a decreasing rate after the closure. Table 4.9 presents the effect of the refinery closure on the housing prices in more details. While the first block of table 4.9 looks at the entire dataset, the second block reduces the sample to municipalities located within a distance of 10 kilometers around the refinery to have a more homogeneous dataset in terms of trends (Currie *et al.*, 2013). While Column 1 does not have any controls, Column 2 adds property characteristics, column 3 includes a municipality fixed effects and the last column adds the unemployment variable.

Figure 4.10: Evolution of the housing prices from 2008 to 2011



Note: This graph represents the evolution of the number of housing prices for the area of Dunkirk versus municipalities without a refinery within the same department. September 2009 corresponds to the refinery closure.

Table 4.9: Reduced form regressions: house

VARIABLES	(1) OLS	(2) OLS	(3) FE	(4) FE	(5) distance<10km OLS	(6) OLS	(7) FE	(8) FE
post_closure	-0.0236	0.0492	0.0556***	0.0505***	0.0144	0.0894**	0.0915**	0.0613**
	(0.0620)	(0.0525)	(0.00517)	(0.00773)	(0.0272)	(0.0239)	(0.0232)	(0.0163)
post	0.0698	0.0451			-0.0335*	0.000681		
	(0.0712)	(0.0663)			(0.0113)	(0.00949)		
closure	0.220***	0.158***			0.161***	0.114***		
	(0.0469)	(0.0389)			(0.0272)	(0.0169)		
pool		0.283***	0.185**	0.168**		0.570***	0.370**	0.278*
		(0.0567)	(0.0603)	(0.0574)		(0.0844)	(0.0978)	(0.0887)
typ_pav		0.209***	0.188***	0.188***		0.174***	0.154***	0.147***
		(0.0146)	(0.0135)	(0.0151)		(0.0158)	(0.0122)	(0.00476)
terrace		0.113***	0.132***	0.131***		0.146	0.123	0.131
		(0.0147)	(0.0115)	(0.0111)		(0.0900)	(0.0840)	(0.0794)
attic		-0.0420**	-0.00213	-0.00499		0.0327	0.0376	0.00522
		(0.0166)	(0.00683)	(0.00583)		(0.0247)	(0.0526)	(0.0371)
balcon		0.140*	0.0691	0.0672		0.0296**	0.0436	0.0550
		(0.0705)	(0.0416)	(0.0456)		(0.00658)	(0.0278)	(0.0254)
srf_parking		-0.000521	-0.000512	-0.000694		-0.00431**	-0.00260	-0.00415*
		(0.000831)	(0.000728)	(0.000725)		(0.000746)	(0.00165)	(0.00159)
house_srf		0.00508***	0.00487***	0.00532***		0.00364*	0.00349**	0.00531***
		(0.000451)	(0.000429)	(0.000129)		(0.00126)	(0.00105)	(0.000188)
less_5_years		0.257***	0.237***	0.236***		0.244***	0.237***	0.253***
		(0.0173)	(0.0165)	(0.0157)		(0.0295)	(0.0226)	(0.0222)
nbr_pieces		0.0331***	0.0334***	0.0282***		0.0427**	0.0394***	0.0252**
		(0.00522)	(0.00534)	(0.00434)		(0.00963)	(0.00656)	(0.00632)
Un				-0.0332				-0.0373
				(0.0201)				(0.0570)
Year FE	x	x	x	x	x	x	x	x
municipality FE			x	x			x	x
distance<10km					x	x	x	x
Observations	13,870	8,745	8,745	8,499	3,066	1,509	1,509	1,263
Adjusted R-squared	0.018	0.408	0.510	0.522	0.046	0.380	0.451	0.485

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of house. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4.9 shows a rise in the price of houses after the closure and it is robust in all fixed effects models even when I present a more homogeneous dataset in the second block. Adding property characteristics and a municipality fixed effects improve the specification of the model. Unemployment, the economic activity proxy added in the last column reduces slightly the effect of the shock on house prices. The 5 micrograms per cubic meter reduction in SO_2 pollution I observe after the refinery closure leads to approximately 6% increase in housing prices.

After the refinery closure, housing prices in the global model significantly increase by 6% in average.

In addition, I test for the heterogeneity of the treatment results by splitting the sample into different price levels.

Table 4.10: Reduced form regressions: houses with respect to prices

VARIABLES FE	price<median OLS FE	(2) FE	(3) FE	price>median	(6) OLS	(7) FE	price>200000 FE	(10)	(11) OLS
post_closure	0.0630*	0.0674**	0.0642**	-0.00825	-0.00380	-0.00893	0.0690**	0.0645***	0.0632***
	(0.0301)	(0.0197)	(0.0217)	(0.0224)	(0.00794)	(0.0108)	(0.0225)	(0.0114)	(0.0126)
post	0.0181			0.0595**			0.0155		
	(0.0396)			(0.0221)			(0.0235)		
treatment	0.0986***			0.0914***			0.0175		
	(0.0237)			(0.0159)			(0.00985)		
house_type	0.143***	0.139***	0.135***	0.0409**	0.0426**	0.0463***	-0.0127	-0.00987	-0.00972
	(0.00622)	(0.00835)	(0.0106)	(0.0124)	(0.0127)	(0.0111)	(0.0114)	(0.0151)	(0.0152)
terrace	0.115***	0.140***	0.140***	0.0327**	0.0410***	0.0426***	0.0293*	0.0383*	0.0387*
	(0.0248)	(0.0204)	(0.0211)	(0.0118)	(0.00813)	(0.00855)	(0.0128)	(0.0190)	(0.0190)
attic	0.00367	0.0276**	0.0258**	-0.0182*	-0.00286	-0.00502	-0.00669	-0.000751	-0.00113
	(0.0151)	(0.00805)	(0.00781)	(0.00862)	(0.0122)	(0.0110)	(0.0121)	(0.0166)	(0.0168)
balcony	0.129**	0.138***	0.143***	0.0722	-0.00295	-0.00744	0.0373	0.0149	0.0164
	(0.0412)	(0.0274)	(0.0278)	(0.0690)	(0.0405)	(0.0423)	(0.0287)	(0.0205)	(0.0199)
parking_srf	0.00167***	0.00114**	0.00106**	-0.00357***	-0.00292**	-0.00324**	0.00121	0.00180	0.00182
	(0.000410)	(0.000396)	(0.000392)	(0.000805)	(0.00102)	(0.00103)	(0.00441)	(0.00443)	(0.00442)
garden							-0.0447***	-0.0606**	-0.0573**
							(0.0126)	(0.0185)	(0.0185)
house_srf	0.00328***	0.00339***	0.00335***	0.00260***	0.00268***	0.00315***	0.00174***	0.00198***	0.00204***
	(0.000254)	(0.000251)	(0.000240)	(0.000447)	(0.000425)	(0.000153)	(0.000188)	(0.000166)	(0.000163)
less_5_years	0.0841	0.118	0.123	0.0750***	0.0896***	0.0916***	0.0378**	0.0417**	0.0433**
	(0.0597)	(0.0679)	(0.0682)	(0.0130)	(0.0127)	(0.0145)	(0.0145)	(0.0149)	(0.0145)
room_nb	0.0368***	0.0343***	0.0353***	0.0109	0.0115*	0.00656*	0.00370	0.00366	0.00310
	(0.00371)	(0.00424)	(0.00411)	(0.00601)	(0.00521)	(0.00323)	(0.00277)	(0.00311)	(0.00318)
Un			-0.0315			-0.0178*			-0.0142
			(0.0170)			(0.00801)			(0.00952)
pool				0.243***	0.211***	0.199***	0.147**	0.0943*	0.0931*
				(0.0583)	(0.0489)	(0.0472)	(0.0616)	(0.0444)	(0.0440)
Year FE	x	x	x	x	x	x	x	x	
municipality FE		x	x		x	x		x	x
Observations	4,323	4,323	4,161	4,422	4,422	4,338	1,764	1,764	1,759
Adjusted R-squared	0.192	0.294	0.296	0.277	0.403	0.425	0.210	0.385	0.390

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of flat by level of flat prices. Robust standard errors in parentheses. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

In table 4.10, prices are set below and above the median of housing prices. The median of housing prices is set up at 141 270 Euros in the dataset. I also create a subsample of expensive houses set above 200 000 Euros. The difference in difference estimator is again positive and significant when I consider subsamples below the median and above 200 000 Euros. However, I do not find any significant effect on the subsample above the median.

After the refinery closure, housing prices below the median of prices and above 200 000 Euros increase significantly by 6% in average.

In view of these estimations, housing buyers in general seem willing to pay for an improvement in air quality.

Flats Prices Table 4.11 and 4.12 focus on the effect of the refinery closure on the price of flats. Models are similar to the one presented for houses. While table 4.11 looks at the entire dataset, table 4.12 splits the sample with respect to flat prices. While environmental quality increases, Table 4.11 shows a reduction in the price of flats after the closure. The result is still robust when I add unemployment as a control in the last column. Note that the significanc is quite low. However, the significance disappears in the second part of table 4.12 when I consider a more homogeneous sample with municipalities not further than 10 kilometers away.

Table 4.11: Reduced form regressions: flat

VARIABLES	(1) OLS	(2) OLS	(3) FE	(4) FE	(5) distance<10km OLS	(6) OLS	(7) FE	(8) FE
post_treatment	-0.120	-0.102**	-0.0771*	-0.0784*	0.0270	-0.0394	-0.0209	-0.0329
	(0.0804)	(0.0304)	(0.0357)	(0.0344)	(0.125)	(0.0822)	(0.0778)	(0.113)
post	0.0417	0.146***			-0.140	0.100		
	(0.0969)	(0.0288)			(0.102)	(0.0619)		
treatment	0.0392	0.0813**			0.0642	0.307**		
	(0.0562)	(0.0234)			(0.0787)	(0.0767)		
nbr_niveau		0.0208***	0.0192***	0.0197***		0.00416	0.00504	0.00587
		(0.00570)	(0.00446)	(0.00460)		(0.00351)	(0.00418)	(0.00434)
flat_type		0.0888	0.0876*	0.0857*		0.147**	0.149**	0.144**
		(0.0586)	(0.0430)	(0.0419)		(0.0381)	(0.0363)	(0.0380)
terrace		0.0931**	0.0346	0.0369		0.0481	0.0503	0.0482
		(0.0351)	(0.0276)	(0.0265)		(0.0345)	(0.0386)	(0.0362)
attic		-0.0599	-0.0538	-0.0519		-0.0552	-0.0597	-0.0562
		(0.0421)	(0.0363)	(0.0359)		(0.0545)	(0.0570)	(0.0559)
balcon		0.146***	0.0691***	0.0671***		0.0348	0.0292	0.0237
		(0.0175)	(0.0183)	(0.0186)		(0.0298)	(0.0322)	(0.0347)
srf_parking		-6.44e-05	-0.000397	0.000141		-0.00162	-0.00204	-0.00207
		(0.00153)	(0.00184)	(0.00210)		(0.00322)	(0.00346)	(0.00338)
garden		0.0578	0.0384	0.0426		0.0775	0.0700	0.0826
		(0.0489)	(0.0489)	(0.0506)		(0.0370)	(0.0356)	(0.0437)
house_srf		0.00838***	0.00993***	0.00991***		0.0103***	0.0104***	0.0104***
		(0.000462)	(0.000469)	(0.000516)		(0.000622)	(0.000602)	(0.000703)
less_5_years		0.485***	0.605***	0.605***		0.529***	0.527***	0.527***
		(0.0815)	(0.0986)	(0.0965)		(0.0680)	(0.0672)	(0.0657)
room_nb		0.0211	0.0223	0.0221		0.0246	0.0254	0.0264
		(0.0129)	(0.0146)	(0.0153)		(0.0147)	(0.0150)	(0.0157)
Un				-0.0704				-0.105**
				(0.0463)				(0.0226)
Year FE	x	x	x	x	x	x	x	x
municipality FE			x	x			x	x
distance<10km					x	x	x	x
Observations	2,848	2,065	2,065	2,026	1,269	875	875	836
Adjusted R-squared	0.006	0.515	0.630	0.629	0.011	0.620	0.625	0.621

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of flat. All regressions are estimated using OLS, with standard errors clustered at the year and department level. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

After the refinery closure, flat prices decrease significantly by 7% in average, without flat prices differentiation.

In addition, I test for the heterogeneity of the treatment results by splitting the sample into different price levels. In table 4.12, prices are set below and above the median for flat prices. I also create a subsample of expensive flats set above 150 000 Euros. The median of flat prices is set up at 118 797 Euros in the dataset. Results are not significant for flats above the median. However, the difference in difference estimator becomes positive and significant in Table 4.12 when I consider a subsample set above 150 000 euros.

Table 4.12: Reduced form regressions: flat with respect to prices

	price<mean			price>mean			price>150000		
VARIABLES	OLS	FE	FE	OLS	FE	FE	OLS	FE	FE
post_closure	-0.00361	-0.0230	-0.0301*	-0.00920	0.0135	0.0137	0.0649	0.0661***	0.0647***
	(0.0293)	(0.0170)	(0.0135)	(0.0217)	(0.0194)	(0.0188)	(0.0352)	(0.0156)	(0.0159)
post	-0.0469*			0.0848**			0.0613		
	(0.0201)			(0.0291)			(0.0464)		
treatment	0.0673**			0.0138			0.00863		
	(0.0212)			(0.0154)			(0.0277)		
level_nb	0.00947	0.0115**	0.0121**	0.0177***	0.0158***	0.0158***	0.0204***	0.0195***	0.0193***
	(0.00510)	(0.00416)	(0.00419)	(0.00276)	(0.00354)	(0.00371)	(0.00437)	(0.00424)	(0.00439)
flat_type	0.178***	0.182***	0.183***	-0.0604	-0.0569	-0.0571	-0.0409*	-0.0587*	-0.0583*
	(0.0272)	(0.0375)	(0.0365)	(0.0386)	(0.0402)	(0.0401)	(0.0198)	(0.0254)	(0.0255)
terrace	0.0861*	0.0478	0.0493	0.0512**	0.0172	0.0173	0.0509	0.0339	0.0337
	(0.0444)	(0.0407)	(0.0401)	(0.0194)	(0.0125)	(0.0126)	(0.0270)	(0.0224)	(0.0225)
attic	-0.0150	-0.0103	-0.00728	-0.151***	-0.145***	-0.145***	-0.186***	-0.188***	-0.190***
	(0.0287)	(0.0261)	(0.0259)	(0.0196)	(0.0199)	(0.0198)	(0.0284)	(0.0246)	(0.0259)
balcony	0.103**	0.0565**	0.0556**	0.0490***	0.0265	0.0266	0.0319	0.00303	0.00258
	(0.0324)	(0.0230)	(0.0228)	(0.0132)	(0.0206)	(0.0206)	(0.0219)	(0.0187)	(0.0191)
parking_srf	0.00844***	0.00416	0.00446	-0.00235	-0.00193	-0.00191	-0.00379	-0.00424	-0.00435
	(0.00221)	(0.00272)	(0.00250)	(0.00186)	(0.00251)	(0.00260)	(0.00341)	(0.00380)	(0.00384)
garden	-0.00912	0.000892	0.00667	0.0592	0.0292	0.0294	0.0277	-0.00931	-0.0108
	(0.0409)	(0.0405)	(0.0412)	(0.0369)	(0.0408)	(0.0407)	(0.0333)	(0.0403)	(0.0401)
house_srf	0.00540***	0.00642***	0.00621***	0.00369***	0.00512***	0.00512***	0.00277***	0.00399***	0.00400***
	(0.000452)	(0.000387)	(0.000438)	(0.000426)	(0.000526)	(0.000526)	(0.000431)	(0.000469)	(0.000473)
less_5_years	0.312***	0.336***	0.334***	0.151***	0.256***	0.257***	0.0614**	0.140***	0.139***
	(0.0403)	(0.0877)	(0.0869)	(0.0222)	(0.0251)	(0.0252)	(0.0211)	(0.0261)	(0.0263)
room_nb	0.0340**	0.0381***	0.0403***	0.00490	0.00618	0.00626	0.00623	0.00732	0.00732
	(0.0101)	(0.00955)	(0.0110)	(0.0101)	(0.0114)	(0.0115)	(0.00919)	(0.00942)	(0.00944)
Un			-0.0712**			-0.00264			0.0173
			(0.0258)			(0.0238)			(0.0299)
Year FE	x	x	x	x	x	x	x	x	
municipality FE		x	x		x	x		x	x
Observations	1,060	1,060	1,024	1,005	1,005	1,002	597	597	597
Adjusted R-squared	0.416	0.501	0.502	0.258	0.384	0.382	0.263	0.371	0.371

Note: This table presents the coefficient estimates of the reduced form estimate of the effect of refineries closure on the log price of flat by level of flat prices. Robust standard errors in parentheses. Statistical significance is denoted by: *** p<0.01, ** p<0.05, * p<0.1

After the refinery closure, prices of expensive flats, above 150 000 Euros, increase significantly by 6% in average.

After the refinery closure, prices of cheap flats, below the median of flat prices, decrease significantly by 3% in average.

While the refinery closure has a positive and significant impact on pollution and on the severity of respiratory outcomes, results from the hedonic price analysis do not always reflect economic benefits of an improvement in air quality. Results vary substantially with respect to the segment being considered. Shedding light on both methods shows first a lack of significant results when the hedonic approach is applied on the overall property market. On the contrary, house prices, increasing after the refinery closure, are in line with the environmental economic intuition, whereas flat prices, decreasing after the refinery closure, do not comply with the theoretical literature. Previous results show that there are no changes in the job market and in the demographic composition of the studied area. Thus, one may wonder why the price of cheap flats tends to decrease after the refinery closure in Dunkirk, although we observe health and air quality improvments.

Table 4.13 shows the distribution of property prices by social categories. The first block of this descriptive table indicates more than 90% of factory workers buy houses relatively to flat whereas 30% of executives buy flats. However, we observe from the second and third blocks, most executives buy expensive flats and houses whereas most factory workers buy cheaper houses and flats. Thus, factory workers, the main social category working in an oil refinery, may be the source of the decline in flat prices. Results from an an independent samples t-test indicate that there is a statistically significant difference between property prices and social class ($p = 0.000$). In other words, executives display a statistically significantly higher flat price (148035.7) than factory workers (96704.46).

Table 4.13: The distribution of flats and houses with respect to the social class of the buyer

	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others	Total
property									
flat	372	504	306	110	150	362	26	36	2848
%	29.41	17.36	14.15	13.87	6.58	37.36	14.77	16.67	17.04
House	893	2,400	1,856	683	2,130	607	150	180	13870
%	70.59	82.64	85.85	86.13	93.42	62.64	85.23	83.33	82.96
Total	1,265	2,904	2,162	793	2,280	969	176	216	16718
%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Flat prices									
Below the median	117	277	208	41	117	105	11	13	1424
%	31.45	54.96	67.97	37.27	78	29.01	42.31	36.11	50
Above the median	255	227	98	69	33	257	15	23	1424
%	68.55	45.04	32.03	62.73	22	70.99	57.69	63.89	50
Total	372	504	306	110	150	362	26	36	2848
%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
House prices									
Below the median	214	1016	961	295	1368	201	62	99	6935
%	23.96	42.33	51.78	43.19	64.23	33.11	41.33	55	50
Above the median	679	1384	895	388	762	406	88	81	6935
%	76.04	57.67	48.22	56.81	35.77	66.89	58.67	45	50
Total	893	2400	1856	683	2130	607	150	180	13870
%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: This table presents the distribution of flat and houses with respect to the socioeconomic category in the number and in % of each social category. It also presents the distribution of flats and houses prices with respect to the median of the price.

The increase/decrease of property prices comes from an increase/decrease of the net demand.[7] If the net demand of cheap flats is mainly represented by factory workers, the decrease of flat prices comes from a decrease of their net demand, meaning factory workers want to leave the area. Table 4.14 details the effect of the closure on each social class category. Thanks to a really precise dataset on each transaction, I am able to present results on the impact of the closure on sellers (block 1) and buyers (block 2) separately by social category. Sample size for social category by buyers and sellers are similar because explicative variables are dummies. It is coded as one if a buyer or a seller belongs to a specific category and 0 if not. Columns contain separated logit regression models. Table 4.14 indicates that after the refinery closure, factory workers among others are significative sellers of their properties in Dunkirk compared to the control group (column 5). On the contrary, executives sell less in Dunkirk after the closure than the control group (column 1). The Intermediate profession group is not significant in both sellers and buyers estimations. Looking at the buyer's estimation, the coefficient for factory workers is not significative. The coefficient for executive is negative suggesting those executives are buying less in Dunkirk after the shock than they are in the other municipalities. It suggests executives may decide not to sell their property in Dunkirk following the closure but the reduction in pollution may not be strong enough to be an incentive for executives to come to live in the area.

7. This is due to an increase/decrease of the household demand who want to live in the area or an increase/decrease of the household demand who want to leave the area

Table 4.14: The effect of the closure on social class

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SELLERS								
Variables	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others
post_closure	-0.0241***	0.00978	0.0309**	-0.00209	0.0228***	0.0377**	0.00418**	0.0174**
	(0.00286)	(0.00949)	(0.0102)	(0.00477)	(0.00635)	(0.0136)	(0.00130)	(0.00535)
Observations	16,718	16,718	16,718	16,718	16,718	16,718	16,718	16,718
Adjusted R-squared	0.026	0.032	0.026	0.034	0.046	0.060	0.036	0.034
BUYERS								
Variables	executive	intermediate professions	employees	craftsman	factory workers	retired	farmers	others
post_treatment	-0.0313**	0.0223	0.0344***	0.00860**	0.0167	-0.00498	-0.00624**	0.00524
	(0.0126)	(0.0134)	(0.00915)	(0.00293)	(0.0122)	(0.0113)	(0.00216)	(0.00410)
Observations	16,718	16,718	16,718	16,718	16,718	16,718	16,718	16,718
Adjusted R-squared	0.034	0.043	0.031	0.036	0.065	0.033	0.040	0.021

Note: This table presents the effect of the closure on the migration by social category. Robust standard errors in parentheses. Statistical significance is denoted by: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Some elements of explanation may be drawn from these results. Unless I assume flat owners are less informed about pollution impact, objectivity from the damage function versus perception from the hedonic price analysis do not explain results differences between houses and flats. The perception of the decrease in health risk may indeed be smaller than the real decrease in health risk because of imperfect information as explained in the theoretical section previously. Besides, because their budget constraint is tight, buyers of cheap flats spending constraint may be stronger than others. In the case of the refinery closure, the shock not only has an impact on pollution but also on economic activity. Even if they were aware of air pollution reduction, the economic effect is larger in relative terms for owners/buyers of cheap flats, set below the median in the subsample of table 4.12, than the environmental effect. While executives, sensitive to the reduction in pollution prefer to stay in Dunkirk, factory workers, sensitive to the economic shock, prefer to sell their property in Dunkirk. On the contrary, the environmental effect is larger in relative terms for owners/buyers of expensive flats and houses than the economic effect. A basic finding in the literature is that income tends to influence willingness to pay for environment positively and significantly (Hokby & Soderqvist, 2003). Willingness to pay for environment may differ with the level of income. In this context, people buying expensive dwellings may be more sensitive to pollution than people buying cheap dwellings.

Behind such complexity in population behavior, the monetary evaluation with the hedonic price analysis, alone, cannot give an absolute value (Maslianskaïa-Pautrel, 2009). Hedonic results can be used by policy makers provided a careful interpretation and/or comparing with other methods such as the evaluation of sanitary costs.

4.5.4. Monetary evaluation

I derive an approximation of the cost of pollution in term of labor by looking at the cost of hospital admission.[8] The ExternE project gives a monetary evaluation that I take into account to derive an approximation of the cost of pollution. In ExternE, the working hypothesis has been to use the exposure response functions for particles and for O3 as a basis. Effects of SO_2 are assumed to arise indirectly from the particulate nature of nitrate and sulfate aerosols, and they are calculated by applying the particle exposure response(ER) functions to these aerosol concentrations. The cost of one hospital admission and one emergency room visit for respiratory illness is evaluated at 2 000 Euros per admission and 670 Euros per visit respectively (price year 2000) (Bickel & Friedrich, 2005). While I do not find any effects on the number of respiratory admissions, I find an effect on the length of stay. The effect size that I find suggests that $5 \mu g/m^3$ decrease in sulfur dioxide pollution subsequent to the closure (Table 4.2) leads to a decrease in the length of stay at the hospital close to one day (Table 4.4). On the opposite, the ExternE project finds an effect on the number of admissions. The ExternE project calculates the annual rate of attributable emergency respiratory hospital admission at 7.03 per 10 $\mu g/m^3$ per

8. The ExternE project (External Costs of Energy), project of the European Commission, aims to measure the damages to society which are not paid for by its main actors and to translate these damages into a monetary value.

100,000 people (all ages). Using ExternE calculation, the $5\mu g/m^3$ decrease in sulfur dioxide pollution leads to 3,5 per 100,000 people emergency respiratory admissions less. Assuming a population of 376 439 in 2009 (INSEE) people in Dunkirk and a cost for respiratory admissions of 2 000 Euros, ExternE would have approximately estimated an environmental economic benefits for respiratory admission of 26,350 Euros in Dunkirk. Nevertheless, the absence of results for the number of admission in our case suggests pollution effects may be non linear. As one day of hospital traditional admission costs 700 Euros according to the high Council for the future of health insurance and the yearly average number of hospital respiratory admission in Dunkirk is 700, estimations from this studdy suggest a yearly cost difference of nearby 490,000 Euros between before and after the refinery closure in Dunkirk.

However, results from other studies suggest that the monetary value of respiratory admissions impacts is not high, compared with mortality from long-term exposure. The part that is quantified in this study (associated with shorter hospital stays) is likely to substantially underestimate the total health benefit because it omits effects on mortality. I consider now the possible cost of chronic mortality using exposure response functions from the ExternE project. The ExternE numbers, used here, are based on the years of life lost (YOLL) per death. Exposure response functions from the ExternE project are used to compute mortality loss. I take the slope of concentration response function [cases/(person year $\mu g/m^3$)] (SCR) slope for chronic mortality for France (all causes, male + female) as: 2.77exp(-4). Assuming a cost for Year of Life Loss of 50,000 per year, the 5 micrograms of pollution costs 12 683 euros per person per year. Assuming a population of 376 439 in 2009 (INSEE) people in Dunkirk, it amounts to 4 774 million a year. There is a yearly cost difference of approximately 4 774 million of Euros a year in terms of chronic mortality between before and after the refinery closure in Dunkirk. Note that these costs are underevaluated due to the monthly average used for the estimation. By aggregating daily data to monthly data, I smooth possible peaks of pollution that can have a strong impact on health. In addition, Individuals' valuations of the lower health risk and any effect on neighborhood aesthetics may be reflected in the price differential associated with proximity to the refinery. The difference between the health benefits and the price differential may be interpreted as a combination of individual valuations of neighborhood aesthetics and people perception of a decrease in health risk.

In comparison, the housing prices increase by 5 percent after the refinery closure which suggests an average benefit of approximately 7,500 Euros for each transaction. Hence the benefits accrue to all house owners as a wealth shock, equal to roughly 5% of the housing value. Note that this gain underestimates the potential benefit for all properties. In fact, it exists as a latent gain on every transaction that has not been realized. The health costs I estimate are above the change in prices I observe but the mortality costs are likely to be much higher. It suggests another interpretation for the discrepancies betweeen these two kinds of results; maybe the housholds are aware and conscious of the morbidity costs but not entirely conscious of the mortality costs.

4.6. Conclusion

This paper tests the short term effect of sulfur dioxide (SO_2) on respiratory outcomes and the global impact it has on housing prices. Our first goal is to assess the impact of air pollution reduction on health outcomes for those municipalities that experienced a reduction in air pollution following the activity stop of the oil refinery in Dunkirk, north of France, in September 2009. Since I have a panel dataset, the best way to isolate the causal effect of the reduction in SO_2 concentration from the closure at one oil refinery is to examine outcome differences between Dunkirk and its counterpart's overtime. I look at the effects of a closure on local measures of sulfur dioxide (SO_2) concentration. I address several longstanding issues dealing with non-random selection and behavioral responses to air pollution that may bias previous studies. This result is particularly significantfor at risk populations such as children below 5 years old and people over 70 years old. I also find a significant effect for adults between 25 and 40 years old suggesting air pollution concentration can have a deleterious impact on labor outcomes.

The second part of the project aims to examine the willingness to pay for an improvement in air quality and the objective effect of a decrease in toxic concentration on respiratory outcomes. To do so, I use a wide and rich dataset on property prices. I find first a positive effect of an improvement in air pollution on housing prices. The environmental effect is larger in relative terms for buyers of houses and expensive flats than the economic effect. Buyers of cheap flats are less sensitive to pollution than others due to budget constraint.

The first results indicate that SO_2, even at levels below current air quality standards in most of the world, has significant negative impacts on the severity of a respiratory disease, suggesting that the strengthening of regulations on SO_2 pollution would yield additional benefits. The second part of the project suggests failing to incorporate different segment into the hedonic approach analysis will lead to biased estimates.

Annex 1

I derive a a measure of expected exposure as a sum of probability of air pollution exposure. And I describe my probability of air pollution exposure as follows: The probability $Pr(m|LOS)$ of being admitted in month m, given the length of stay (LOS), is calculated as follows:

$$Pr(m|LOS) = \begin{cases} Pr(m|LOS = LOS_0) = P(m|m) \text{ if LOS} \leq 30 \\ Pr(m|LOS = LOS_1) = Pr(m-1|m) \text{ if LOS } \in [30;60] \\ Pr(m|LOS = LOS_2) = Pr(m-2|m) \text{ if LOS } \in [60;90] \\ Pr(m|LOS = LOS_3) = Pr(m-3|m) \text{ if LOS } \in [90;120] \\ ... \\ Pr(m|LOS = LOS_n) = Pr(m-(n)|m) \text{ if LOS } \in [990;1020] \end{cases}$$

$Pr(m|LOS)$ represents the probability of being exposed to air pollution (or being admitted) in month m given the length of stay (LOS). $P(m|m)$ is the probability a patient has been exposed in month m, given m the same month of discharge, if he has stayed up to 30 days. $Pr(m-1|m)$ represents the probability a patient has been exposed in month m-1, given m the same month of discharge, if he has stayed between 30 to 60 days.

The same way, the probability $Pr(m-1|LOS)$ of being admitted in month m-1, given the length of stay (LOS), is calculated as follows:

$$Pr(m-1|LOS) = \begin{cases} Pr(m-1|LOS = LOS_0) = Pr(m-1|m) \text{ if LOS} \leq 30 \\ Pr(m-1|LOS = LOS_1) = Pr(m-2|m) \text{ if LOS } \in [30; 60] \\ Pr(m-1|LOS = LOS_2) = Pr(m-3|m) \text{ if LOS } \in [60; 90] \\ Pr(m-1|LOS = LOS_3) = Pr(m-4|m) \text{ if LOS } \in [90; 120] \\ ... \\ Pr(m-1|LOS = LOS_n) = Pr(m-(n+1)|m) \text{ if LOS} \in [990; 1020] \end{cases}$$

$Pr(m-1|m)$ represents the probability a patient has been admitted in month m-1 given the length of stay up to 30 days. $Pr(m-2|m)$ represents the probability a patient has been admitted in month m-2 given the length of stay up from 30 to 60 days, etc...

The maximum length of stay is 1020 days in the dataset.

Conclusion générale

L'analyse de l'impact de la pollution sur la santé et des disparités sociales reste encore aujourd'hui un enjeu important en économie de l'environnement.

Les effets sanitaires de beaucoup de polluants sont établis dans la littérature. Non seulement ces effets varient selon leurs concentrations et combinaisons mais aussi selon l'exposition et la vulnérabilité des populations. Dans ce contexte, les risques sanitaires sont difficiles à quantifier. Il n'en reste pas moins qu'il existe un lien entre statut socioéconomique et le dioxyde d'azote (NO_2) auquel nous nous intéressons dans le premier chapitre qui lie la mortalité en France aux disparités environnementales. Les résultats confirment l'existence de relations de long terme entre la pollution de fond et la mortalité et soulèvent également des questions importantes à propos de la justice environnementale en France. Les immissions de polluants étudiées se situent en dessous des seuils de pollution fixés par les autorités publiques, confortant l'idée que la pollution est fluctuante et que même des niveaux faibles de concentration peuvent être actifs.

Le deuxième chapitre apporte un éclairage original sur l'impact du dioxyde de soufre (SO_2) sur la santé des enfants à la naissance. Les résultats sont dans la lignée d'Hanna & Oliva (2011) et montrent que la fermeture d'une raffinerie entraîne bien une diminution de la pollution avec un impact positif sur la santé des enfants à la naissance. Le dioxyde de soufre, même à des niveaux de pollution en dessous des seuils fixés par les autorités publiques, continue d'affecter les personnes fragiles, asthmatiques, détenteurs de problèmes pulmonaires. Les résultats corroborent l'idée qu'un renforcement de la régulation concernant le dioxyde de soufre SO_2 serait bénéfique pour la société et particulièrement pour les populations vivant proches des raffineries. Il reste important de ne pas sous-estimer les impacts sanitaires des polluants atmosphériques même à de très faibles niveaux de concentrations.

L'objectif du troisième chapitre est l'évaluation des conséquences de la pollution atmosphérique suite à la fermeture définitive de la raffinerie des Flandres à Dunkerque. L'approche par les prix hédoniques et par la fonction de dommage sont mises en regard pour évaluer respectivement la perception subjective des agents et les bénéfices sanitaires objectifs. Selon le segment considéré, la perception des agents sur la qualité de l'air ne semble pas toujours converger avec l'existence d'effets bénéfiques d'une amélioration de la qualité de l'air sur la santé. Malgré des effets sanitaires positifs, les classes peu aisées ne sont pas sensibles à une amélioration de la qualité de l'air. Ce chapitre soulève des questions sur les préférences des agents économiques à propos d'un choc qui se veut à la fois environnemental et économique. La théorie économique par l'effet revenu ou par la considération de biens supérieurs suggère que les riches préfèrent une société avec moins de dégradation environnementale que les pauvres. Les études empiriques abondent dans le même sens. Il semble pourtant que les classes défavorisées sont plus exposées à la pollution, véritable facteur d'engagement environnemental. Par ailleurs, les pures préférences des agents ne sont révélées qu'en cas d'information parfaite et de dépenses contraintes similaires. Or, l'approche segmentaire retenue dans cette thèse montre qu'il est nécessaire d'étudier les hétérogénéités à l'intérieur de chaque groupe socioéconomique afin de mettre en place des politiques visant à réduire les nuisances environnementales dans une perspective de réduction des inégalités. La contrainte budgétaire semble en effet plus serrée pour les propriétaires d'appartement, par exemple, qui auraient des dépenses contraintes plus importantes par rapport à leur budget que les autres.

Bibliography

Adair, Alastair, Berry, Jim, & McGreal, Stanley. 1994. Investment and private sector residential development in inner-city Dublin. *Journal of Property Valuation and Investment*, **12**(1), 47–56.

Airparif. 2013. Plan régional de la qualité de l'air.

Almond, Douglas, & Currie, Janet. 2011. Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, **25**(3), 153–72.

Almond, Douglas, Chay, Kenneth Young, & Lee, David S. 2002. Does Low Birth Weight Matter?: Evidence from the US Population of Twin Births.

Almond, Douglas, Edlund, Lena, & Palme, Marten. 2009. Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. *The Quarterly Journal of Economics*, **124**(4), 1729–1772.

Annesi-Maesano, Isabella, Moreau, David, Caillaud, Denis, Lavaud, François, Le Moullec, Yvon, Taytard, André, Pauli, Gabrielle, & Charpin, Denis. 2007. Residential proximity fine particles related to allergic sensitisation and asthma in primary school children. *Respiratory medicine*, **101**(8), 1721–1729.

Bajari, Patrick, Fruehwirth, Jane Cooley, il Kim, Kyoo, & Timmins, Christopher. 2012. A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution. *American Economic Review*, **102**(5), 1898–1926.

Baumol, William J, & Oates, Wallace E. 1988. *The theory of environmental policy.* Cambridge University Press.

Beatty, Timothy K.M., & Shimshack, Jay P. 2011. School buses, diesel emissions, and respiratory health. *Journal of Health Economics*, **30**(5), 987 – 999.

Besson, Danielle. 2006. *Consommation de tabac : la baisse s'est accentuée depuis 2003.* Tech. rept.

Bickel, Peter, & Friedrich, Rainer. 2005. *ExternE, Externalities of Energy, Methodology 2005 update.* Tech. rept.

Bishai, D., & Opuni, M. 2009. Are infant mortality rate declines exponential? The general pattern of 20th century infant mortality rate decline. *Popul Health Metr*, **7**.

Black, Sandra E, Devereux, Paul J, & Salvanes, Kjell G. 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, **122**(1), 409–439.

Blanchard, Olivier J, & Summers, Lawrence H. 1986. Hysteresis and the European unemployment problem. *NBER Macroeconomics Annual*, **1**, 15–90.

Bolte, Gabriele, Tamburlini, Giorgio, & Kohlhuber, Martina. 2010. Environmental inequalities among children in Europe evaluation of scientific evidence and policy implications. *The European Journal of Public Health*, **20**(1), 14–20.

Brookshire, David S, Thayer, Mark A, Schulze, William D, & d Arge, Ralph C. 1982. Valuing public goods: a comparison of survey and hedonic approaches. *The American Economic Review*, **72**(1), 165–177.

Butler, Richard V. 1982. The specification of hedonic indexes for urban housing. *Land Economics*, **58**(1), 96–108.

Cambois, Emmanuelle, & Jusot, Florence. 2011. Contribution of lifelong adverse experiences to social health inequalities: findings from a population survey in France. *The European Journal of Public Health*, **21**(5), 667–673.

Campagna, Dave, Lefranc, Agnes, Nunes-Odasso, Catherine, & Ruth, Ferry. 2003. Evaluation des risques de la pollution urbaine sur la santé en Ile-de-France (erpurs) : liens avec la mortalité 1987-1998. *Observatoire régional de santé d'Ile-de-France.*

Chanel, Olivier, Scapecchi, Pascale, & Vergnaud, Jean-Christophe. 2000. Pollution atmosphérique et santé: évaluation monetaire et effets de long terme. *Université Panthéon-Sorbonne (Paris 1).*

Chay, Kenneth, Dobkin, Carlos, & Greenstone, Michael. 2003. The Clean Air Act of 1970 and Adult Mortality. *Journal of Risk and Uncertainty*, **27**(3), 279–300.

Chay, Kenneth Y., & Greenstone, Michael. 2003. The Impact Of Air Pollution

On Infant Mortality: Evidence From Geographic Variation In Pollution Shocks Induced By A Recession. *The Quarterly Journal of Economics*, **118**(3), 1121–1167.

Chay, Kenneth Y., & Greenstone, Michael. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, **113**(2), 376–424.

Clougherty, Jane E. 2010. A growing role for gender analysis in air pollution epidemiology. *Environmental Health Perspectives*, **118**(2), 167.

Cropper, Maureen L. 2000. Has Economic Research Answered the Needs of Environmental Policy? *Journal of Environmental Economics and Management*, **39**(3), 328–350.

Cunningham, FG, Leveno, KJ, Bloom, SL, Hauth, JC, Gilstrap III, LC, & Wenstrom, KD. 2005. Fetal growth and development. *Williams Obstetrics*, **23**.

Currie, Janet, & Neidell, Matthew. 2004. *Air Pollution and Infant Health: What Can We Learn From California's Recent Experience*. NBER Working Papers 10251. National Bureau of Economic Research, Inc.

Currie, Janet, & Walker, Reed. 2011. Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, **3**(1), 65–90.

Currie, Janet, Hanushek, Eric A, Kahn, E. Megan, Neidell, Matthew, & Rivkin, Steven G. 2009. Does Pollution Increase School Absences? *The Review of Economics and Statistics*, **91**(4), 682–694.

Currie, Janet, Greenstone, Michael, & Moretti, Enrico. 2011. Superfund Cleanups and Infant Health. *American Economic Review*, **101**(3), 435–41.

Currie, Janet, Davis, Lucas, Greenstone, Michael, & Walker, Reed. 2013. Do Housing Prices Reflect Environmental Health Risks? Evidence from More than 1600 Toxic Plant Openings and Closings. *NBER Working Papers*, **18700**.

Davis, Lucas W. 2004. The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review*, **94**(5), 1693–1704.

Deguen, Severine, & Zmirou, Denis. 2012. Social inequalities resulting from health risks related to ambient air quality; A European review. *European Journal of Public Health*, **20**(1), 27–35.

Delucchi, Mark, Murphy, James, & McCubbin, Donald. 2002. *The Health and Visibility Cost of Air Pollution: A Comparison of Estimation Methods*. Tech. rept.

Deschenes, Olivier, Greenstone, Michael, & Guryan, Jonathan. 2009. Climate Change and Birth Weight. *American Economic Review*, **99**(2), 211–17.

Environmental Protection Agency (EPA). 2011. *Sulfur Dioxide @ONLINE*.

Environmental Protection Agency (EPA). 2012a. *Control Emmissions Technologies - Transport Dispersion of Air Pollutants @ONLINE*.

Environmental Protection Agency (EPA). 2012b. *Sulfur Dioxide @ONLINE*.

Ettner, Susan L. 1996. New evidence on the relationship between income and health. *Journal of Health Economics*, **15**(1), 67–85.

Feitelson, Eran I, Hurd, Robert E, & Mudge, Richard R. 1996. The impact of airport noise on willingness to pay for residences. *Transportation Research Part D: Transport and Environment*, **1**(1), 1–14.

Finkelstein, Murray M, Jerrett, Michael, DeLuca, Patrick, Finkelstein, Norm, Verma, Dave K, Chapman, Kenneth, & Sears, Malcolm R. 2003. Relation between income, air pollution and mortality: a cohort study. *Canadian Medical Association Journal*, **169**(5), 397–402.

Fiscella, Kevin, & Franks, Peter. 2000. Individual income, income inequality, health, and mortality: what are the relationships? *Health services research*, **35**, 307.

Fletcher, Mike, Gallimore, Paul, & Mangan, Jean. 2000. Heteroscedasticity in hedonic house price models. *Journal of Property Research*, **17**(2), 93–108.

Frankenberg, Elizabeth, McKee, Douglas, & Thomas, Duncan. 2005. Health consequences of forest fires in Indonesia. *Demography*, **42**(1), 109–129.

Gillette, Donald G. 1977. *Oxidant Air Pollution and Work Performance of Citrus Harvest Labor.*

Graff Zivin, Joshua, & Neidell, Matthew. 2012. The Impact of Pollution on Worker Productivity. *American Economic Review*, **102**(7), 3652–73.

Graff Zivin, Joshua, & Neidell, Matthew. 2013. Environment, Health, and Human Capital. *Mimeo, Columbia University.*

Greenstone, Michael, & Gallagher, Justin. 2008. Does Hazardous Waste Matter?

Evidence from the Housing Market and the Superfund Program. *The Quarterly Journal of Economics*, **123**(3), 951–1003.

Greenstone, Michael, & Hanna, Rema. 2011. *Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.* Working Papers 1114. Massachusetts Institute of Technology, Center for Energy and Environmental Policy Research.

Griliches, Zvi, & Mason, William M. 1972. Education, Income, and Ability. *Journal of Political Economy*, **80**(3), S74–S103.

Grineski, Sara, Bolin, Bob, & Boone, Christopher. 2007. Criteria Air Pollution and Marginalized Populations: Environmental Inequality in Metropolitan Phoenix, Arizona. *Social Science Quarterly*, **88**(2), 535–554.

Groot, Wim, & Maassen van den Brink, Henriette. 2007. The health effects of education. *Economics of Education Review*, **26**(2), 186–200.

Grossman, Michael. 2000. The human capital model. *Handbook of health economics*, **1**, 347–408.

Gunasekara, Fiona Imlach, Carter, Kristie, & Blakely, Tony. 2011. Change in income and change in self-rated health: Systematic review of studies using repeated measures to control for confounding bias. *Social Science & Medicine*, **72**(2), 193–201.

Hanna, Brid Gleeson. 2007. House values, incomes, and industrial pollution. *Journal of Environmental Economics and Management*, **54**(1), 100–112.

Hanna, Rema, & Oliva, Paulina. 2011. *The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City*. NBER Working Papers 17302. National Bureau of Economic Research.

Harper, Krista, Steger, Tamara, & Filcak, Richard. 2009. Environmental justice and Roma communities in Central and Eastern Europe. *Environmental Policy and Governance*, **19**(4), 251–268.

Hoechle, Daniel. 2006 (Nov.). *XTSCC: Stata module to calculate robust standard errors for panels with cross-sectional dependence*. Statistical Software Components, Boston College Department of Economics.

Hokby, Stina, & Soderqvist, Tore. 2003. Elasticities of Demand and Willingness to Pay for Environmental Services in Sweden. *Environmental & Resource Economics*, **26**(3), 361–383.

INSEE. 2010. *L'inscription territoriale de la raffinerie des Flandres dans la région Nord-Pas-de-Calais*. Tech. rept.

Janke, Katharina, Propper, Carol, & Henderson, John. 2009. Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics*, **18**(9), 1031–1055.

Joyce, Theodore, Grossman Michael, & Goldman, Fred. 1988a. An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health. *Journal of Urban Economics*, **25**(2), 32–51.

Joyce, Theodore, Grossman Michael, & Goldman, Fred. 1988b. An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health. *Journal of Urban Economics*, **25**(2), 32–51.

Katsouyanni, Klea, Zmirou, D, Spix, C, Sunyer, J, Schouten, JP, Ponka, A, Anderson, HR, Le Moullec, Y, Wojtyniak, B, Vigotti, MA, *et al.* 1995. Short-term effects of air pollution on health: a European approach using epidemiological time-series data. The APHEA project: background, objectives, design. *European respiratory journal*, **8**(6), 1030–1038.

Krupnick, Alan, Alberini, Anna, Cropper, M, Simon, N, O'Brien, B, Goeree, R, & Herntzelman, M. 2002. Age, Health and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents. *Journal of Risk and Uncertainty*, **24**(2), 161–86.

Kuminoff, Nicolai V., Parmeter, Christopher F., & Pope, Jaren C. 2010. Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management*, **60**(3), 145–160.

Kunzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., Herry, M., Horak, F., V., Puybonnieux-Texier Jr, & Quenel, P. 2000. Public-Health Impact of Outdoor and Traffic-related Air Pollution: A European Assessment. *Lancet*, **356**, 795–801.

Laigle, L, & Oehler, V. 2004. Les enjeux sociaux et environnementaux du développement urbain: la question des inégalités écologiques. *Final Report, Centre Scientifique et Technique du Bâtiment, Paris. CSTB.*

Lancaster, KJ. 1966. A new approach to consumer theory. *Journal of Political Economy*, **74**, 132– 157.

Laurent, Eloi. 2011. Issues in environmental justice within the European Union. *Ecological Economics*, **70**(11), 1846–1853.

Laurent, Olivier, Bard, Denis, Filleul, Laurent, & Segala, Claire. 2007. Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. *Journal of epidemiology and community health*, **61**(8), 665–675.

Lecoq, Pierre, Dor, Frédéric, & Kairo, Cécile. 2009. *Description des valeurs repères toxicologiques utilisées lors d'expositions aiguës par inhalation des populations*. Report. The French Institute for Public Health Surveillance, InVS.

Lindahl, Mikael. 2005. Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human Resources*, **40**(1), 144–168.

Maantay, Juliana. 2007. Asthma and air pollution in the Bronx: methodological and data considerations in using GIS for environmental justice and health research. *Health & place*, **13**(1), 32.

Marr, Linsey C. Matthew E. R. 2010. Effect of Air Pollution on Marathon Running Performance. *Medicine Science in Sports Exercise*, **42**(3), 585.

Martens, Willem JM. 1998. Climate change, thermal stress and mortality changes. *Social science & medicine*, **46**(3), 331–344.

Maslianskaïa-Pautrel, Masha. 2009. La valorisation de la qualité de l'air par l'approche hédonique: une revue de la littérature. *Revue française d'économie*, **23**(3), 109–160.

McLeod, H., Langford, I. H., Jones, A. P., Stedman, J. R., Day, R. J., Lorenzoni, I., & Bateman, I. J. 2000. The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. *Regional Environmental Change*, **1**, 78–85.

Midorikawa, Kaoru, Uchida, Takafumi, Okamoto, Yoshinori, Toda, Chitose, Sakai, Yoshie, Ueda, Koji, Hiraku, Yusuke, Murata, Mariko, Kawanishi, Shosuke, & Kojima, Nakao. 2004. Metabolic activation of carcinogenic ethylbenzene leads to oxidative DNA damage. *Chemico-biological interactions*, **150**(3), 271–281.

Mohai, Paul, Lantz, Paula M, Morenoff, Jeffrey, House, James S, & Mero, Richard P. 2009. Racial and socioeconomic disparities in residential proximity to polluting industrial facilities: evidence from the Americans' Changing Lives Study. *Journal Information*, **99**(S3).

Morello-Frosch, Rachel, Pastor, Manuel, & Sadd, James. 2002. Integrating environmental justice and the precautionary principle in research and policy making: the case of ambient air toxics exposures and health risks among schoolchildren in

Los Angeles. *The annals of the American academy of political and social science*, **584**(1), 47–68.

Moretti, Enrico, & Neidell, Matthew. 2011. Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. *Journal of Human Resources*, **46**(1), 154–175.

Neidell, Matthew. 2004. Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, **23**(6), 1209–1236.

Neidell, Matthew. 2009. Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations. *Journal of Human Resources*, **44**(2).

Orford, Scott. 2004. Identifying and comparing changes in the spatial concentrations of urban poverty and affluence: a case study of inner London. *Computers, Environment and Urban Systems*, **28**(6), 701–717.

Palmquist, Raymond B. 1984. Estimating the Demand for the Characteristics of Housing. *The Review of Economics and Statistics*, **66**(3), 394–404.

Parry, Ian, & Small, Kenneth. 2002. *Does Britain or the United States Have the Right Gasoline Tax?* Discussion Papers dp-02-12-. Resources For the Future.

Parry, Ian W. H., & Small, Kenneth A. 2005. Does Britain or the United States Have the Right Gasoline Tax? *American Economic Review*, **95**(4), 1276–1289.

Pascal, Laurence, Blanchard, Myriam, Fabre, Pascal, Larrieu, Sophie, Borrelli, David, Host, Sabine, Chardon, Benoit, Chatignoux, Edouard, Prouvost, Hélène, Jusot, Jean-François, *et al.* 2009. *Liens à court terme entre la mortalité et les admissions à l'hôpital et les niveaux de pollution atmosphérique dans neuf villes françaises.* Tech. rept. 5.

Pastor, Manuel, Sadd, Jim, & Hipp, John. 2001. Which Came First? Toxic Facilities, Minority Move-In, and Environmental Justice. *Journal of Urban Affairs*, **23**(1), 1–21.

Pearce, Jamie R, Richardson, Elizabeth A, Mitchell, Richard J, & Shortt, Niamh K. 2010. Environmental justice and health: the implications of the socio-spatial distribution of multiple environmental deprivation for health inequalities in the United Kingdom. *Transactions of the Institute of British Geographers*, **35**(4), 522–539.

Peng, Roger, Welty, Leah, & McDermott, Aidan. 2004 (July). *The National Morbidity, Mortality, and Air Pollution Study Database in R.* Johns Hopkins University Dept. of Biostatistics Working Paper Series 1044. Berkeley Electronic Press.

Portney, Paul R. 1981. Housing prices, health effects, and valuing reductions in risk of death. *Journal of Environmental Economics and Management*, **8**(1), 72 – 78.

Premji, S., Bertrand, F., Smargiassi, A., & Daniels, M. 2007. Socioeconomic Correlates of Municipal Level Pollution Emissions on Montreal Island. *Canadian Journal of Public Health*, **98**(2), 138–42.

RABL, Ari. 2003. Mortalité due à la pollution de l'air: comment interpréter les résultats. *Pollution atmosphérique*, **45**(180), 541–550.

Rady, Sven. 1998. *Boom In, Bust Out: Young Households and the Housing Price Cycle*. FMG Discussion Papers dp310. Financial Markets Group.

Ransom, Michael R, & Iii, C. Arden Pope. 1995. External Health Costs Of A Steel Mill. *Contemporary Economic Policy*, **13**(2), 86–97.

Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, **82**(1), 34–55.

Sanders, Nicholas J., & Stoecker, Charles F. 2011 (Sept.). *Where Have All the Young Men Gone? Using Gender Ratios to Measure Fetal Death Rates*. NBER Working Papers 17434. National Bureau of Economic Research, Inc.

Schikowski, T., Sugiri, D., Reimann, V., Pesch, B., Ranft, U., & Ursula, K. 2008. Air Pollution Exposure in Urban Areas to Social Differences in Respiratory Health. *BMC public health*, 8–179.

Schulz, Laura C. 2010. The Dutch Hunger Winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences*, **107**(39), 16757–16758.

Scruggs, Lyle A. 1998. Political and economic inequality and the environment. *Ecological Economics*, **26**(3), 259–275.

Soleille, Sébastien. 2004. *L'industrie du raffinage et le devenir des fiouls lourds*. Report. French national institute for industrial environment and risks, INERIS.

Stein, TP., Schluter, MD., Soteropoulos, PA., Galante, AT., & Tolias, PP. 2003. Scholl TO Fetal lymphocyte gene expression and pregnancy outcome. *Faseb Journal*, **17**(5).

Stringhini, Silvia, Berkman, Lisa, Dugravot, Aline, Ferrie, Jane E., Marmot, Michael, Kivimaki, Mika, & Singh-Manoux, Archana. 2012. Socioeconomic Status, Structural and Functional Measures of Social Support, and Mortality: The British Whitehall II Cohort Study. *American Journal of Epidemiology*.

Subramanian, SV, & Kawachi, Ichiro. 2006. Being well and doing well: on the importance of income for health. *International Journal of Social Welfare*, **15**(s1), S13–S22.

ud Din, Zia, Emmett, Pauline, Steer, Colin, & Emond, Alan. 2013. Growth Outcomes of Weight Faltering in Infancy in ALSPAC. *Pediatrics*, **131**(3), e843–e849.

Viel, Jean-François, Hägi, Mathieu, Upegui, Erika, & Laurian, Lucie. 2011. Environmental justice in a French industrial region: Are polluting industrial facilities equally distributed? *Health & place*, **17**(1), 257–262.

Wilhelm, Michelle, Qian, Lei, & Ritz, Beate. 2009. Outdoor air pollution, family and neighborhood environment, and asthma in La Fans children. *Health Place*, **15**(1), 25 – 36.

World Health Organization (WHO). 2005. *Effects of air pollution on children's health and development.* Report.

World Health Organization (WHO). 2006. Lignes directrices OMS relatives à la qualité de l'air : particules, ozone, dioxyde d'azote et dioxyde de soufre.

World Health Organization (WHO). 2009. La pollution de l'air à l'interieur des habitations. *Technical report.*

World Health Organization (WHO). 2011. *Air quality and health @ONLINE.*

World Health Organization (WHO). 2012. Environmental health inequalities in Europe. *Technical report.*

Yamada, E., Vanna, A.T., Naspitz, C.K., & Sole, D. 2002. International Study of Asthma and Allergies in Childhood (ISAAC): validation of the written questionnaire (eczema component) and prevalence of atopic eczema among Brazilian children. *J Investig Allergol Clin Immunol*, **12**(1), 34–41.

Yu, Tak-sun Ignatius, Wong, Tze Wai, Wang, Xiao Rong, Song, Hong, Wong, Siu Lan, & Tang, Jin Ling. 2001. Adverse effects of low-level air pollution on the respiratory health of schoolchildren in Hong Kong. *Journal of occupational and environmental medicine*, **43**(4), 310–316.

Zabel, Jeffrey E., & Kiel, Katherine A. 2000. Estimating the Demand for Air Quality in Four U.S. Cities. *Land Economics*, **76**(2), 174–194.

Zivin, Joshua Graff, & Neidell, Matthew. 2013. *Environment, Health, and Human*

Capital. NBER Working Papers 18935. National Bureau of Economic Research, Inc.