

Drut, Marion; Mahieu, Aurélie

**Conference Paper**

## Correcting agglomeration economies: How air pollution matters

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Drut, Marion; Mahieu, Aurélie (2015) : Correcting agglomeration economies: How air pollution matters, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124668>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Correcting agglomeration economies: How air pollution matters

Marion Drut\*, Aurélie Mahieux†

February 11, 2015

## Abstract

This paper aims to correct standard measures of agglomeration economies to account for air pollution generated by commuting. This paper examines the impact of nitrogen oxide ( $\text{NO}_x$ ) a pollutant mainly released by transportation on worker productivity. Literature on agglomeration economies highlights the positive role of employment density on productivity, without accounting for the environmental impact of a better accessibility. First, standard estimates of agglomeration economies for the 304 French employment areas are in line with the literature. Then, we introduce  $\text{NO}_x$  emissions, which reveals that emissions reduce the expected agglomeration gains by more than 13%, confirming that air pollution matters.

**Keywords:** agglomeration economies, accessibility, atmospheric pollution, transportation policies

**JEL:** O18, R23

---

\*EQUIPPE, Université Lille 1, Faculté d'Economie et de Sciences Sociales, Bâtiment SH2, 59655 Villeneuve d'Ascq Cedex, France. Email address: marion.drut@ed.univ-lille1.fr.

†EQUIPPE, Université Lille 1, Faculté d'Economie et de Sciences Sociales, Bâtiment SH2, 59655 Villeneuve d'Ascq Cedex, France. Email address: aurelie.mahieux@ed.univ-lille1.fr.

‡We would like to thank Alain Ayong Le Kama for his precious insights, as well as Moez Kilani for his helpful comments. We thank also the participants of the Mobil.TUM 2014 and ASRDLF 2014 conferences for their productive suggestions. We are also grateful to Quentin David, Clément Nedoncelle and Fabrice Gilles for their useful observations.

# 1 Introduction

Agglomeration economies play a key role in urban economics. The very existence of cities or of any concentration of activities can only be explained in the light of increasing returns in production activities, provided that we rule out the role played by the attributes of physical geography (Fujita and Thisse, 2002). Agglomeration economies are positive externalities derived from the spatial concentration of economic activity (firms and households) that affects the productivity of firms. They are increasing external returns to scale with respect to the size or density of population or employment.

Studies generally estimate the net agglomeration effects and support that agglomeration positively impacts labor productivity. Concentration of economic activity was first defined by the size of the population or employment, then with measures of density. Ciccone and Hall (1996) are the first to propose a framework investigating the effects of employment density on labor productivity. In more recent years, new geography economists such as Combes et al. (2008, 2011) have enhanced the basic framework by adding new elements such as market potential, land area, firms specialization and economic diversity.

Other authors (Graham, 2007; Rice et al., 2006) focus on the effects of a new transportation infrastructure on labor productivity. They conclude that a new infrastructure has a positive effect on accessibility, thus enlarging the opportunities offered to workers and leading to increased labor productivity. Nevertheless, none of the above mentioned studies takes into account the environmental impact generated by an increased accessibility, namely commuting. New transportation infrastructures in particular and enhanced accessibility in general reduce the generalized cost of travel. According to the law of supply and demand, the demand for this good increases, therefore leading to induced traffic (Downs, 1992; Hills, 1996; Cervero, 2002; Noland and Lem, 2002). If the new transportation infrastructure or policy enhancing accessibility does not result in more sustainable mobility patterns, the additional commuting trips will generate higher levels of polluting emissions (Goodwin, 1999; Litman, 2011). In particular, nitrogen oxide ( $\text{NO}_x$ ) emissions primarily result from transportation. Epidemiological studies show that

atmospheric pollution has a negative and significant impact on human health (see e.g., Currie et al. (2009a, 2009b)). The deterioration of health implies both lower labor supply (Ostro, 1983; Hanna and Oliva, 2011; Carson et al., 2011) and lower labor productivity (Lavy et al., 2012; Graff Zivin and Neidell, 2012).

This article aims at correcting estimations of agglomeration economies by accounting for air pollution resulting from commuting. We add air pollution variables in the general framework studying agglomeration economies. More specifically, we explore the impact of  $\text{NO}_x$  emissions on productivity.  $\text{NO}_x$  emissions originate mainly from diesel vehicle exhaust. The objective of the present paper is to show that pollution has to be included in the estimations of agglomeration effects. The results obtained confirm a negative and significant impact of air pollution on productivity.

We use aggregate data for the year 2009 for the 304 French metropolitan employment areas. The employment area level constitutes the relevant spatial unit for transportation projects and policies, as well as for studies related to the labor market (Combes and Lafourcade, 2012). However, very few studies are conducted on such a fine geographic level. In this article, we combine standard data concerning the main determinants of agglomeration economies, such as employment and wages, as well as data on emissions for one air pollution variable,  $\text{NO}_x$ . Data are disaggregated at the industry level into five sectors and then these data are pooled.

First, we estimate the effects on labor productivity per worker of employment density, accessibility measured as a market potential *à la* Harris (1954), surface area, economic diversity, and sectoral specialization. In line with the literature, the results show an increase in productivity of 0.03% for a 1% increase in employment density. Second, we introduce the variable measuring air pollution:  $\text{NO}_x$  emissions. In our specification, we use  $\text{NO}_x$  emissions as a proxy for atmospheric pollution. In line with epidemiological studies, we find that air pollution negatively impacts labor productivity. A 1% increase in the level of  $\text{NO}_x$  emissions leads to almost 0.1% decrease in productivity. Third, we compare the models with and without air pollution. One may expect *prima facie* an increase of the effect of density on productivity when the density variable is “cleaned” from pollution, which is considered as a diseconomy. However, the variation of the effect

of density on productivity depends not only on the effect of pollution on productivity (which is expected to be negative) but also on the correlation between these two variables (which is also negative). As a result, these two negative effects induce a positive bias. Therefore, when pollution is accounted for, the positive effect of employment density on productivity is reduced. Finally, we focus on an illustrative case to show the magnitude of the reduction of agglomeration economies when local air pollution is considered. When  $\text{NO}_x$  emissions are included in the model, the productivity gains of agglomeration are reduced by more than 13%.

Agglomeration economies are often enhanced by new transportation policies or infrastructures that improve accessibility and contribute to the densification of the area. However, improved accessibility induces traffic and therefore pollution emissions. So far as we know, the impact of air pollution on productivity is never addressed in specifications estimating agglomeration effects. In a sustainable development context, these results shed a new light for the assessment of transportation projects such as tramways or Bus with a High Level of Service. This study allows putting into perspective the agglomeration benefits resulting from the implementation of a new transportation infrastructure or policy.

The paper is organized as follows. Section 2 provides a brief review of the literature on agglomeration economies. Section 3 presents the data and descriptive statistics. Section 4 estimates the general econometric model and addresses common endogeneity issues. In Section 5, we introduce the environmental variable and present the adjusted results. In Section 6, we compare both specifications and develop the illustrative case. In Section 7, we draw conclusions.

## **2 Theoretical background of agglomeration economies**

### **2.1 Sources and classification of agglomeration economies**

Already long ago, Marshall (1890) set the assumption that geographic concentration of activities generates productivity gains. Duranton and Puga (2004) explore the theoretical

microeconomic foundations of agglomeration economies. They emphasize three distinct mechanisms leading to agglomeration economies: learning, matching, and sharing. First, learning effects or technological spillovers relate to the generation, the diffusion, and the accumulation of knowledge. The process of learning occurs at small spatial scales, since it requires close interactions and physical proximity. Therefore, dense areas make a higher degree of specialization possible (Ciccone and Hall, 1996). Second, large and dense labor markets allow for better employees/employers matching with lower search costs. Third, large and dense markets lower access costs to both customers and suppliers of intermediate goods and services, even when transportation costs are low (Krugman, 1991). Moreover, this last mechanism allows for the sharing of local public goods and of any other indivisible facilities, as well as the sharing of risks.

A further distinction can be made between “localization economies” and “urbanization economies” (Krugman, 1991; Rosenthal and Strange, 2004), though their sources are similar. Localization economies, also called within-industry externalities, imply increasing returns to scale that are external to the firm but internal to the industry (e.g., technological spillovers, intermediate inputs sharing, labor market matching). Urbanization economies, also called between-industry externalities, refer to agglomeration benefits that are external to the firm or the industry but internal to the city (e.g., local public goods sharing, input-output sharing). In this work, we do not aim at estimating these two kinds of effects separately. Indeed and as stated by Graham (2007), “an aggregate estimate of density externalities is sufficient to demonstrate the relationship between agglomeration, productivity, and transport investment”.

The creation and growth of cities result from two opposing forces: agglomeration (centripetal forces) and dispersion (centrifugal forces) (Krugman, 1991; Fujita and Thisse, 2002). It is usually agreed that agglomeration effects follow a bell-shaped curve (Henderson, 1974; Fujita et al., 1999). Agglomeration economies first exceed diseconomies up to a certain threshold, and lead to concentration of activities. Thereafter concentration of activities leads to congestion and pollution issues, rising land rents, higher labor costs, crime and socio-economic polarization, which constitutes costs for society, and hence a dispersion force. In the literature, these two effects are rarely identified separately, and

only the net effect is usually estimated, as in this study.

## 2.2 Magnitude of agglomeration effects

Several literature reviews are available on this topic (see e.g., Rosenthal and Strange (2004); Puga (2010); Melo et al. (2009)). Although they are drawn on different methodologies and on countries (mainly the US and Europe) of various size and industry-structure, all the studies support evidence that agglomeration economies positively impact labor productivity. Depending on the measure applied, elasticity coefficients for productivity usually range from 0.03 to 0.08 (Rosenthal and Strange, 2004). This means that a 1% increase in either density or city size results in a 0.03 to 0.08% increase in labor productivity. Ciccone and Hall (1996) find that doubling employment density raises the average labor productivity by 6%, and that more than half of the variance in output per worker across US states can be explained by differences in employment density. Ciccone (2002) finds similar results (4.5-5%) for five European countries. Combes et al. (2008, 2011), using the same measure, estimate an elasticity of productivity of about 0.08 on French departments, and of 0.06 on French employment areas with aggregate data, along with an estimate of 0.03-0.04 on French employment areas with individual data. Rice et al. (2006) stress on the fact that studies based on individual data show smaller coefficient values.

## 2.3 The impact of transport

Other authors focus on the effects of a new transportation infrastructure on labor productivity and employment growth. First, “by driving down travel costs, extra roads increase the attractiveness of a city, which brings new residents” and therefore increases employment (Duranton and Turner, 2012). Duranton and Turner (2012) find that a 10% increase in a city’s stock of highways causes a 1.5% increase in its employment. Furthermore, assumption is made that new or improved transportation infrastructures enhance accessibility, which in turn enlarges the concentration of activities from which agglomeration economies arise (Gibbons and Overman, 2009). Venables (2007) explores the theoretical foundations behind the effects of transportation infrastructures on productiv-

ity. He concludes that better accessibility leads to increased productivity. In an empirical study, Rice et al. (2006) and then Matas et al. (2013) confirm this finding and evidence a 1.2% increase in productivity when travel times are reduced by 10%. However, there is evidence of a steep decrease of agglomeration economies with distance (Rice et al., 2006; Graham et al., 2009; Matas et al., 2013). Therefore, a new transportation infrastructure mainly benefits to the surrounding area.

Agglomeration economies are additional benefits that are more and more accounted for in transportation project appraisals as “wider economic benefits” (Vickerman, 2007; DfT, 2005; Victoria Department of Transport, 2012). Additional benefits can be substantial, as reveals the 25% increase in benefits for the London CrossRail project<sup>1</sup> (DfT, 2005). Nevertheless, none of the above mentioned studies takes into account the environmental impact generated by an increased accessibility, namely commuting. Correcting the assessment of agglomeration economies brings new perspectives on transportation project appraisals and allows for a better allocation of public funds.

### 3 Data and descriptive statistics

“A fine level of geographical details” is required to obtain accurate estimates (Ciccone, 2002). For this purpose, we choose to draw our analysis at the employment area level. So far, very few studies have investigated the effects of agglomeration at the employment area level (see Combes et al. (2008, 2010)). Most studies use larger spatial units, such as NUTS 3 areas<sup>2</sup> (Ciccone, 2002; Rice et al., 2006; and Combes et al., 2011). However, the choice of the size (or number) and shape (or the drawing of boundaries) of the spatial units only slightly influences the results, as demonstrated by Briant et al. (2010). French employment areas were defined in 1983 and modified several times thereafter (1994, 1999 and 2010). They are smaller than NUTS 3 areas (French “Departments”), but larger than LAU 1 areas<sup>3</sup> (French “Cantons”). Furthermore and contrary to NUTS or LAU areas, their borders are defined by commuting patterns rather than being administratively

---

<sup>1</sup>The CrossRail project in London is an underground east-west rail link connecting existing rail networks on each side of the city (DfT, 2005).

<sup>2</sup>NUTS stands for Nomenclature of Territorial Units for Statistics.

<sup>3</sup>LAU stands for Local Administrative Unit.



stated. It is admitted that at least 75% of the labor force lives and works within the same employment area. Most employment areas correspond to a metropolitan area or to a city and its catchment area (see Figure 3 in Appendix A). Thus, analyzing the effects of transportation infrastructure on employment areas seems all the more relevant, since employment areas are built on commuting trips. Moreover, small spatial units such as employment areas constitute the appropriate spatial level for studying productivity issues since it has been demonstrated that agglomeration effects decrease rapidly with distance and mainly arise within 80 km.

In 2010, Metropolitan France includes 304 employment areas. We use cross-sectional data for the year 2009, which are aggregated at the employment area level. We combine data from General Census of Population with data on employment and wages for the year 2009. All data are derived from INSEE (French Institute of Statistics and Economic Studies). They are disaggregated at the industry level into five sectors (agriculture, manufacturing, construction, trade and services, public administration), and then pooled. The database is a two-dimension panel: employment area and industry. It consists of 1,520 observations. We use workplace-based data on wages<sup>4</sup> to approximate labor productivity. To obtain employment densities, we use data on the number of jobs<sup>5</sup> divided by the surface areas. Surface areas are in square kilometers. The variable ‘specialization’ is constructed with the employment share of each sector in total area. The measure ranges from 0 when nobody works in a specific sector to 1 when the total employment of the area is concentrated in this sector. We use as a measure of diversity the inverse of Herfindhal Index, applying data on sectoral employment. The measure equals 1 when jobs are concentrated in one sector, 5 when they are perfectly divided into the 5 sectors considered. The market potential of a zone is the sum of the opportunities derived from all the other zones while considering the distance between this zone and all the other ones. An opportunity is defined as the employment density of a particular zone divided by the distance from this zone to another zone. Since French employment areas are built on commuting patterns, it can be assumed that employment centers are usually located at the centroid of the area. Since it constitutes a more accurate measure of accessibility

---

<sup>4</sup>File “Rémunérations” from INSEE.

<sup>5</sup>File “Postes” from INSEE.

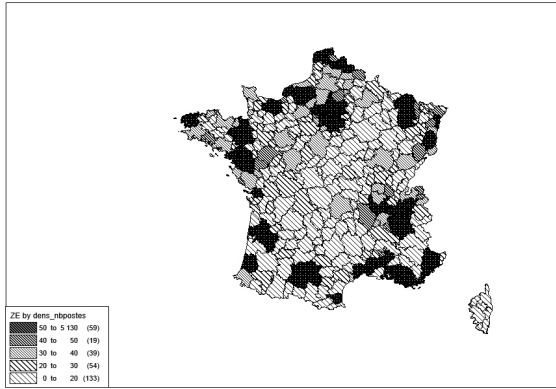


Figure 1: Employment density in French employment areas

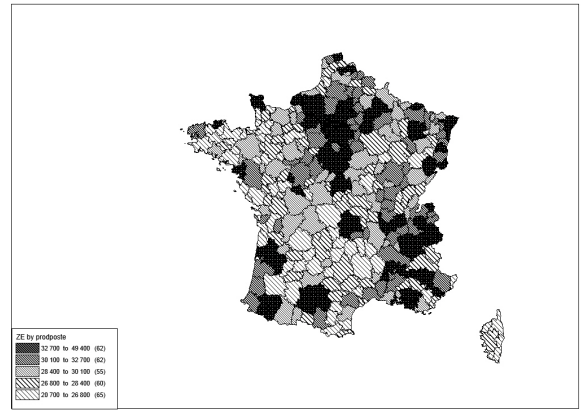


Figure 2: Worker productivity in French employment areas

than Euclidean distance, we compute real road network distances with a Geographical Information System<sup>6</sup> to build the market potential variable.

Table 1: Summary statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Productivity	1,520	24,869.65	4,454.30	11,988.95	49,399.54
Density	1,520	65.66	315.16	2.48	5,124.87
Area	1,520	1,796.87	1,390.35	119.40	8,752.00
Specialization	1,520	0.20	0.15	0.0002	0.64
Diversity	1,520	3.23	0.33	2.09	4.29
Market Potential	1,520	83.67	57.51	25.22	480.53

Figure 1 and Figure 2 show that the highest employment density areas correspond to the most productive areas. These figures illustrate the underlying intuition behind agglomeration economies: labor productivity is likely to be correlated with employment density.

<sup>6</sup>Distances are computed using calcdist-280.mbx tool on MapInfo. With this software, we compute distances between the centroids of each French employment areas.

## 4 The standard model

This section estimates the net effect of employment density on labor productivity per worker. We develop the general framework for French employment areas and control for common endogeneity issues.

### 4.1 The general framework

The basic framework has recently been enhanced by additional explanatory variables measuring urbanization economies, such as accessibility measured as a market potential, surface area, and economic diversity. Sectoral specialization is often added to identify localization economies. Variables used in the general econometric specification are described below.

#### 4.1.1 Common variables used in the literature

In the literature, we observe two main approaches measuring labor productivity. First, productivity can be estimated with the help of a production function using data on value added, since agglomeration economies lead to increased total factor productivity (Rosenthal and Strange, 2004). Second, wage equations are commonly in use to approximate productivity, assuming that at the competitive equilibrium workers receive wages equal to their marginal labor productivity. Rice et al. (2006) show the existence of a strong correlation (0.76) between these two kinds of productivity variables, namely gross value added per employee per hour worked and average hourly earnings. Moreover, the authors stress the fact that for small areas measuring productivity with gross value added may be biased by the spatial allocation of non-wage incomes.

Various measures of concentration are found. Some authors focus on employment, population or industry size (Sveikauskas, 1975; Segal, 1976; Henderson, 1986) or working age population size (Rice et al., 2006), while others apply measures of density. Ciccone and Hall (1996) define density as ‘the intensity of labor, human, and physical capital relative to physical space’. They are the first to propose a framework investigating the effects of employment density on labor productivity. Density is a continuous variable that

is far less sensitive to the geographic boundaries used than measures of size.

When people and goods are mobile, employment areas are interconnected by migration and trade flows. These interactions have an influence on labor productivity (Head and Mayer, 2004, 2006). In the literature, two families of accessibility measures are in use: effective density and market potential (Matas et al., 2013). The effective density, as applied by Graham (2007) and Matas et al. (2013), is a comprehensive measure of both the accessibility to activity concentration within a specific area and from this area to the other areas. The market potential, derived from Harris (1954) and applied by Combes et al. (2008, 2011), measures only the accessibility to activity concentration of a particular area to the other areas<sup>7</sup>. For this reason, in any specification the market potential has to be used jointly with a measure of the size or density for each area. It is worth noting that changes in transportation infrastructure or policy modify the market potential of a particular area since the relative proximities of activity are altered.

The surface of employment areas is added in order to distinguish density effects from pure scale effects. Indeed, surfaces vary significantly between areas and can impact density effects. Moreover, it is common to introduce a diversity index to capture the local distribution of jobs between the various economic sectors, as well as a measure of sectoral specialization to indicate the within-industry concentration.

#### **4.1.2 Formalizing the standard model**

Following Combes et al. (2008, 2011), this article uses the employment density as a measure of concentration, and the average wage per worker as dependent variable. As prescribed by Moretti (2004), we use nominal wages. In this article, we use the market potential variable, since it best allows for discriminating between the effect of density and the effect of accessibility. Finally, we add other common variables, namely the surface of employment areas, a diversity index and a measure of sectoral specialization.

---

<sup>7</sup>A limit of the market potential measure is that accessibility to foreign countries is not accounted for. This may bias coefficient estimates of border areas.

The general specification is the following:

$$\ln prod_{zs} = \alpha + \beta \ln dens_z + \rho \ln MP_z + \delta \ln area_z + \eta \ln div_z + \theta \ln spe_{zs} + \gamma_s + \varepsilon_{zs},$$

where  $prod_{zs}$  is the average labor productivity per worker for sector  $s$  in zone  $z$ ,  $dens_z$  the employment density in zone  $z$ ,  $MP_z$  the market potential of zone  $z$ ,  $area_z$  the surface of employment area  $z$ ,  $div_z$  a measure of the economic diversity of zone  $z$ ,  $spe_{zs}$  the average sectoral specialization of zone  $z$ ,  $\gamma_s$  the industry fixed effects, and  $\varepsilon_{zs}$  the error term. All variables are measured at the employment area level. In line with the recent literature, we use logs of the variables. The coefficient estimates are then interpreted as elasticities with respect to the different variables.

Table 2 shows the correlation between all variables. As expected, the variable ‘productivity’ is clearly and positively correlated with the variable ‘density’. Table 2 also indicates that the specialization of the area is a factor contributing to higher productivity. In addition, results reveal that ‘density’ and ‘accessibility’ are strongly correlated. ‘Specialization’, ‘density’ and ‘market potential’ seem to have a positive correlation with labor productivity. Employment area surface and diversity are negatively correlated with labor productivity.

Table 2: Correlation matrix

Variables	$\ln prod$	$\ln dens$	$\ln area$	$\ln spe$	$\ln div$	$\ln MP$
$\ln prod$	1.0000					
$\ln dens$	0.3401	1.0000				
$\ln area$	-0.0146	-0.3192	1.0000			
$\ln spe$	0.3505	-0.0962	0.0268	1.0000		
$\ln div$	-0.2181	-0.4059	-0.0078	0.1176	1.0000	
$\ln MP$	0.2089	0.4244	-0.3144	-0.0452	-0.0435	1.0000

Table 3 presents estimation results for robust Ordinary Least Squares (OLS) in the general framework. Variables are introduced successively according to the importance of their correlation with productivity. In line with the literature, we find an elasticity of productivity with respect to density of 0.05. All estimated variables are significant

at the 1% level. Market potential is positive and highly significant too. Its magnitude is comparable to that of density. Both specialization of a zone and its surface impact positively on labor productivity. As found by Combes et al. (2008), the coefficient for economic diversity is negative. Indeed, in the existing literature, there is no consensus on the effect of economic diversity, which can be either negative or positive. Positive effects of economic diversity are generally well-known (Jacobs, 1969). Nevertheless, there also exists negative effects. For instance, the larger the number of sectors in a zone, the more intense the competition to access intermediary goods within this zone. In this case, subcontractors tend to scatter, which confirms the negative effect that diversity can have. Therefore, a negative sign for the coefficient of the diversity variable may mean that negative effects prevail over positive ones.

Table 3: Estimation results for robust Ordinary Least Squares (OLS)

Variables	OLS1	OLS2	OLS3	OLS4	OLS5
$\ln spe$	0.0448*** (0.004)	0.0495*** (0.004)	0.0495*** (0.004)	0.0496*** (0.004)	0.0509*** (0.003)
$\ln dens$		0.0638*** (0.004)	0.0580*** (0.005)	0.0629*** (0.005)	0.0517*** (0.005)
$\ln MP$			0.0286** (0.009)	0.0385*** (0.009)	0.0447*** (0.009)
$\ln area$				0.0294*** (0.005)	0.0254*** (0.005)
$\ln div$					-0.2166*** (0.048)
N	1,520	1,520	1,520	1,520	1,520
$R^2$	0.123	0.264	0.269	0.283	0.298

Note: Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

## 4.2 Controlling for endogeneity issues

Endogeneity issues are then controlled with instruments commonly in use in the literature. Unbiased results are finally presented.

### 4.2.1 Common instruments used in the literature

The OLS method assumes that the explanatory variables are uncorrelated with the error term. Otherwise, coefficient estimates are biased. However, two potential sources of endogeneity are identified in standard econometric specifications related to agglomeration economies: simultaneity bias and omitted variable bias. Simultaneity bias, also called reverse causality, arises when either firms or workers migrate to locations with high productivity, leading therefore to higher densities. Graham et al. (2010) analyze the direction of causality between productivity and agglomeration. They find substantial evidence of reverse causality, in particular for localization economies. This bias would lead to a 20% overestimation of agglomeration economies (Combes and Lafourcade, 2012; Combes et al., 2008, 2011). Omitted variable bias, or unobserved heterogeneity, is particular features impacting productivity but which are not explicitly accounted for in the specification. For instance, the industry mix of a zone or specific geographic characteristics (e.g., climate or relief) may impact productivity (Combes et al., 2010). Factor endowments such as public goods or natural resources play as well a role in determining productivity levels. The level of education of workers is also a leading determinant for wages (Ciccone and Hall, 1996; Combes et al., 2011). Agglomeration effects can be either over or underestimated when variables are omitted.

Combes and Lafourcade (2012) provide a literature review of the solutions usually implemented to correct these biases. The most common approach to deal with the simultaneity bias is to use long lags on population size or population density as instrumental variables (Ciccone and Hall, 1996; Rice et al., 2006; Combes et al., 2008, 2010, 2011). The underlying assumption is that previous patterns of population concentration are correlated with current population or employment densities (the endogenous variable), but are independent from current labor productivity.

Furthermore, firm selection issues may also lead to biased agglomeration effects. Firm selection refers to the fact that large and dense markets are more competitive and hence exclude less productive firms. Therefore, higher productivity in larger or denser areas is the result of a selection process, where only the more productive firms survived. However, Combes et al. (2012) reveal that firm selection is not an important bias for agglomeration

economies estimates.

#### 4.2.2 Instrumenting endogenous variables in the standard model

Since both density and market potential are likely to be endogenous, we instrument both variables. We first instrument employment density using NUTS 3 population densities from 1866 and 1891. We then instrument market potential using NUTS 3 population density from 1866 over inter-zones distances as a measure. Then, unobservable heterogeneity can be controlled for by introducing fixed effects (Glaeser and Maré, 2001). In this study, we use industry fixed effects to control for sectoral heterogeneity.

Table 4 shows the results of various estimations of standard agglomeration economies. Introducing industry fixed-effects slightly modifies the coefficients. Moreover, industry fixed-effects raise the  $R^2$  significantly. Instrumenting potentially endogenous variables leads to a slight decrease in the density coefficient, from 0.050 to 0.027. The results are in line with the literature when education is not accounted for<sup>8</sup>. We also observe that the magnitude and significance of market potential decrease after addressing endogeneity issues.

The Stock and Yogo critical values for the Cragg-Donald F-Statistic are 13.43 for 10% maximum IV bias. The endogeneity C-stat confirms that instrumentation is needed for density and market potential. According to the Cragg-Donald F-stat and Kleibergen-Paap statistic, instruments are not weak. The Hansen J-stat shows that the set of instruments is exogenous.

Finally, given the spatial nature of the study, we check the spatial autocorrelation by computing the Moran's Index for each variable of the model. For this purpose, we build a rook weights matrix, i.e., a contiguity-based matrix in which contiguity is defined by shared borders.<sup>9</sup> The  $p$ -values for the Moran's I statistic indicate for each variable that

---

<sup>8</sup>As highlighted in Combes et al. (2011), introducing the human capital decreases significantly the magnitude of density effects for recent periods. Ciccone and Hall (1996) and Combes and Lafourcade (2012) also warn against the existence of a sorting effect. Highly-skilled workers tend to concentrate in densely populated areas, and they get accordingly higher wages. Variables related to workers' education must be added to the specification in order to control for heterogeneity of skills among workers. However as this paper aims at correcting 'standard' estimates of agglomeration economies with pollution features, we prefer to keep the specification as standard as possible.

<sup>9</sup>The  $p$ -values for the Moran's I of each variable are indicated in brackets: labor productivity of worker (0.86), employment density (0.94), market potential (0.92), diversity (0.52), specialization (0.54), and



Table 4: Standard agglomeration economies: results for various estimation methods

Variables	OLS5	OLS6	IV1	IV2
$\ln dens$	0.0517*** (0.005)	0.0504*** (0.004)	0.0271*** (0.007)	0.0272*** (0.007)
$\ln MP$	0.0447*** (0.009)	0.0435*** (0.007)	0.0555*** (0.008)	0.0554*** (0.008)
$\ln area$	0.0254*** (0.005)	0.0257*** (0.004)	0.0182*** (0.004)	0.0182*** (0.004)
$\ln div$	-0.2166*** (0.048)	-0.1895*** (0.042)	-0.2715*** (0.044)	-0.2711*** (0.044)
$\ln spe$	0.0509*** (0.003)	0.0290** (0.010)	0.0249* (0.010)	0.0249* (0.010)
Industry fixed-effects	No	Yes	Yes	Yes
N	1,520	1,520	1,520	1,520
$R^2$	0.298	0.592	-	-
Cragg-Donald F-stat	-	-	383.224	383.224
Kleibergen-Paap Statistic	-	-	300.934	300.934
Hansen J-Stat	-	-	0.002	0.002
Chi-sq P-value	-	-	0.9631	0.9631
Endogeneity C-stat	-	-	32.327	32.327
Chi-sq P-value	-	-	0.000	0.000

Note: OLS5: No fixed-effects; OLS6: Industry fixed-effects; IV1: Generalized Method of Moments (GMM); IV2: Two Step Least Squares (2SLS); IV1 and IV2: we use log of NUTS 3 population density from 1866 and 1891 to instrument the variable ‘ $\ln dens$ ’. Variable ‘ $\ln MP$ ’ is instrumented by the log of the market potential with population density from 1866. Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

we cannot reject the null hypothesis of no spatial autocorrelation. Therefore, there is no need to use spatial econometric models.

## 5 The extended model: including $\text{NO}_x$ emissions

First, we expose the effect of local air pollution, and especially  $\text{NO}_x$ , on the human health and on labor productivity. Then, we develop the extended specification where a pollution variable ( $\text{NO}_x$  emissions) is added to the standard framework of agglomeration area (0.82).

economies.

## 5.1 The effect of pollution on health and productivity

The link between pollution and health has first been assessed through epidemiological studies on mortality rates. For instance, Lave and Seskin (1970) measure the long-term effects of sulfur oxides and particulates on mortality rates. Then, studies have been carried out on the effects of pollution on morbidity, focusing on variations in labor supply. Ostro (1983) demonstrates that a 10% increase in particulate levels generates a 4.4% decrease in work loss days. Carson et al. (2011) evidence a 8% decrease in household labor supply in Bangladesh due to arsenic exposure. Hanna and Oliva (2011) show that a 1% increase in sulfur dioxide results in a 0.61% decrease in the hours worked in Mexico City. These studies generally use hospital outcomes such as length of stay, emergency room visits, or work loss days to measure the impact of several pollutants on health. However, air pollution may affect not only the extensive margin, but also the intensive margin, that is labor productivity. Graff Zivin and Neidell (2012) first demonstrate the impact of ozone pollution on the productivity of agricultural workers in California. Ozone pollution diminishes lung functioning and negatively impacts productivity in physical work, even when the labor supply remains unchanged. Suglia et al. (2008) show that children living near higher levels of fine particulates perform worse on cognitive tests. Similarly, Lavy et al. (2012) find a negative relationship between both fine particulate matter and carbon monoxide and cognitive performance during school tests. They show that altered cognitive performance results in mis-ranking of students. This may result in inefficient allocation of workers across occupations, and negatively affect labor productivity, especially for intellectual work. In this sense, environmental protection is considered as an investment in human capital sustaining labor productivity and therefore economic growth (Graff Zivin and Neidell, 2012).

In this study, we focus on nitrogen oxide ( $\text{NO}_x$ ). Nitrogen oxide ( $\text{NO}_x$ ) is made of nitric oxide (NO) and nitrogen dioxide ( $\text{NO}_2$ ).  $\text{NO}_2$  is highly toxic and penetrates into the lungs, therefore causing respiratory diseases. NO irritates bronchi and diminishes the oxygen power of blood.  $\text{NO}_x$  emissions result mainly from transport (61%, among which

93% from road transport) due to the exhaust of diesel vehicles. Latza et al. (2009) provide a review of some experimental and epidemiological studies.  $\text{NO}_2$  emissions lead to ear, nose and throat infections, otitis media, respiratory infections and in the most extreme cases myocardial infarctions. In addition, Ghosh et al. (2012) demonstrate an association between  $\text{NO}_x$  and respiratory illnesses (bronchitis and upper airway inflammation) even for levels of  $\text{NO}_x$  lower than the current European Commission standards, especially among very young children.

Although  $\text{NO}_x$  emissions are on a decreasing trend (-45% in France over the period 1990-2011) (CITEPA, 2013), their actual level remains harmful for health. Furthermore, this pollutant affects the environment.  $\text{NO}_x$  are among air pollutants causing acid rains. They also contribute to ozone pollution and to climate change. Although environmental effects are not accounted for in our specification, they are relevant and could be integrated in future analysis.

## 5.2 The extended specification

First estimations of the extended specification are presented, before we control for endogeneity issues and present unbiased results.

### 5.2.1 First estimations

In this article, we use data on  $\text{NO}_x$  emissions for the year 2009 at the NUTS 2 level (French “regions”). Emissions are obtained from each regional AASQA (Association Agréée de Surveillance de la Qualité de l’Air, which is the French regional association for air quality monitoring). The year 2009 is the only available dataset for  $\text{NO}_x$  emissions. Since the specification is defined at an aggregated level, we apply emissions that are an aggregated measure of concentrations recorded at each particular monitoring station. We are aware of the fact that air quality affecting human health is best approximated by concentration levels of pollutants. The relation between concentrations and emissions is complex. For a given level of emissions, concentrations vary depending on meteorological and physical factors such as wind, temperature, humidity, precipitation, topography and height of buildings. In order to partly avoid such bias, we use spatial units which are

much larger than employment areas. Indeed, larger units would better account for wind effects. We obtained pollution data for 21 of the 22 French regions. The following results are therefore drawn on a slightly smaller number of observations than the standard model presented above. We have now 1,485 observations for 297 employment areas.

The extended specification is based on the general framework presented in Section 4.1 and includes the pollution variable for a zone  $z$ , noted ‘ $poll_z$ ’:

$$\ln prod_{zs} = \alpha + \beta \ln dens_z + \rho \ln MP_z + \delta \ln area_z + \eta \ln div_z + \theta \ln spe_{zs} + \lambda \ln poll_z + \gamma_s + \varepsilon_{zs}$$

We test the impact of  $NO_X$  emissions per worker on labor productivity. We integrate the air pollution variable in the general model. Since Lavy et al. (2012) find that pollution has a non-linear impact on productivity, we use the logarithmic form.

Table 5 represents the correlation matrix between all the variables of the general framework and the  $NO_X$  emissions variable. Since correlations between standard agglomeration economies variables are quite similar, complete correlation matrix is not presented in this section. As expected, the correlation matrix shows that  $NO_X$  is negatively correlated with labor productivity.

Table 5: Correlation matrix for  $NO_X$  emissions

Variables	$\ln NO_X$
$\ln prod$	-0.2316
$\ln dens$	-0.3255
$\ln area$	0.2491
$\ln spe$	0.0581
$\ln div$	0.2686
$\ln MP$	-0.3568
$\ln NO_X$	1.0000

Table 6 presents the effect of  $NO_X$  emissions on labor productivity.  $NO_X$  emissions by worker have a negative and significant effect at the 1% level on labor productivity. The

results show that a 1% increase in NO<sub>x</sub> emissions lowers labor productivity by almost 0.07%. Table 6 also indicates that the positive effect of density on productivity is reduced when the variable is instrumented, while the positive effect of accessibility is strengthened. The coefficients of the other variables only slightly differ after instrumentation.

Table 6: The effect of air pollution on productivity

	OLS1	OLS2	IV1	IV2
<i>ln dens</i>	0.0526*** (0.005)	0.0514*** (0.004)	0.0265*** (0.007)	0.0265*** (0.007)
<i>ln MP</i>	0.0374*** (0.009)	0.0365*** (0.006)	0.0452*** (0.007)	0.0452*** (0.007)
<i>ln area</i>	0.0328*** (0.005)	0.0329*** (0.004)	0.0242*** (0.004)	0.0241*** (0.004)
<i>ln div</i>	-0.1456** (0.051)	-0.1229** (0.044)	-0.2137*** (0.047)	-0.2131*** (0.047)
<i>ln spe</i>	0.0515*** (0.003)	0.0323** (0.010)	0.0279** (0.010)	0.0279** (0.010)
<i>ln NO<sub>x</sub></i>	-0.0602*** (0.012)	-0.0595*** (0.008)	-0.0655*** (0.008)	-0.0654*** (0.008)
Industry fixed-effects	No	Yes	Yes	Yes
N	1,485	1,485	1,485	1,485
<i>R</i> <sup>2</sup>	0.318	0.617	-	-
Cragg-Donald F-stat	-	-	359.202	359.202
Kleibergen-Paap Statistic	-	-	300.707	300.707
Hansen J-stat	-	-	0.006	0.006
Chi-sq P-value	-	-	0.9378	0.9378
Endogeneity C-stat	-	-	32.266	32.266
Chi-sq P-value	-	-	0.0000	0.0000

Note: OLS1: No fixed-effects; OLS2: Industry fixed-effects; IV1: Generalized Method of Moments (GMM); IV2: Two Step Least Squares (2SLS); IV1 and IV2: we use log of NUTS 3 population density from 1866 and 1891 to instrument the variable ‘*ln dens*’. Variable ‘*ln MP*’ is instrumented by the log of the market potential with population density from 1866. The Stock and Yogo critical values for the Cragg-Donald F-Statistic are 13.43 for 10% maximum IV bias. As remarked in section 4.2, instrumentation is needed because of endogeneity problems. Besides, the set of instruments is not weak and is exogenous. Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

### 5.2.2 Controlling for endogeneity issues

We are aware of the potential endogeneity bias affecting the pollution variable (reverse causality). On the one hand, the literature introduced in Section 5.1 highlights the causal link between pollution and productivity: pollution impacts negatively on labor productivity. On the other hand, productive regions are likely to pollute more. Therefore, the causal link between pollution and productivity may be reversed.

Previous results constitute first estimations of the effect of air pollution on productivity. They could be enhanced with instrumental variables, such as car ownership rates. We expect  $\text{NO}_x$  emissions to be positively correlated with car ownership rates. Generally, high levels of car ownership rates mean higher car availability, and therefore more trips carried out by car, resulting in higher levels of air pollution. In addition, car ownership rates may also be correlated with productivity, since higher wages facilitate access to cars. Nevertheless, car ownership patterns change rapidly overtime, and we expect lagged car ownership rates not to be correlated with present wages. We use car ownership rates from 1999 as instrument for pollution emissions.

Table 7 presents results for the extended specification when the endogeneity of the pollution variable is controlled. The results slightly differ from the first estimations presented in Section 5.2.1. The density coefficient is reduced from 0.0265 to 0.0253, which indicates that the positive effect of density on productivity is lowered when the endogeneity of the pollution variable is controlled. In addition, the  $\text{NO}_x$  emissions coefficient decreases from -0.0655 to -0.1031, which indicates a stronger negative effect of pollution on productivity. The impact of air pollution on labor productivity remains negative and highly significant, with a 1% increase in air pollution leading to a 0.1% decrease in productivity. According to the standard tests on instrumented variables, the set of instruments used is valid.

In addition, we test the interaction between  $\text{NO}_x$  emissions and density. The interaction term (-0.0186) is negative and significant at the 5% level (see Table 9 in Appendix B), which is in line with the results of the literature on local air pollutants. Consequently,  $\text{NO}_x$  emissions negatively impact the effect of density on productivity. The denser an

area, the more polluted it is, and the more acute health problems will be. Indeed, health problems directly impact workers' productivity, as demonstrated in the literature.

Finally, other endogeneity issues can be suspected between the pollution variable and the other explanatory variables. In this respect, we provide in Appendix C the correlation matrix of coefficients of the regress model. As indicated in Table 10, the correlations between the estimated coefficients are low and confirm the fact that there is no multicollinearity problem in the extended model presented in this section.

## 6 How air pollution reduces agglomeration gains

This section draws a comparison between the two econometric models and estimates the extent to which expected agglomeration gains are reduced when air pollution is accounted for in agglomeration economies estimates.

### 6.1 Comparing the two econometric models

Agglomeration gains are revealed by the elasticity of productivity with respect to density. Estimating the magnitude of the correction of the agglomeration economies requires the comparison between the density coefficients of both models, namely the standard model and the extended model. For this purpose, identical samples are needed. Table 8 presents the results of the standard model on the same sample as the extended model presented in Table 7.

When pollution is accounted for, the density coefficient decreases from 0.0287 to 0.0253, which clearly highlights a reduction in the positive effect of density on productivity. Due to the negative correlation between the density variable and the pollution variable, and the negative sign of the pollution coefficient, we obtain a positive bias, meaning that the standard model overestimates the positive effect of density on productivity (Wooldridge, 2009). A 1% increase in density now leads to a 0.025% increase in labor productivity, instead of the standard 0.029% increase in productivity. Agglomeration economies are therefore reduced when pollution is introduced in the model.

Table 7: The effect of air pollution on productivity after controlling for endogeneity biases

	IV1	IV3
$\ln dens$	0.0265*** (0.007)	0.0253*** (0.007)
$\ln MP$	0.0452*** (0.007)	0.0373*** (0.009)
$\ln area$	0.0242*** (0.004)	0.0266*** (0.004)
$\ln div$	-0.2137*** (0.047)	-0.1881*** (0.047)
$\ln spe$	0.0279** (0.010)	0.0285** (0.010)
$\ln NO_X$	-0.0655*** (0.008)	-0.1031*** (0.021)
Industry fixed-effects	Yes	Yes
N	1,485	1,485
Cragg-Donald F-stat	359.202	81.291
Kleibergen-Paap Statistic	300.707	51.135
Hansen J-stat	0.006	0.116
Chi-sq P-value	0.9378	0.7334
Endogeneity C-stat	32.266	32.878
Chi-sq P-value	0.0000	0.0000

Note: IV1: Generalized Method of Moments (GMM); IV3: Generalized Method of Moments (GMM); IV1 and IV3: we use log of NUTS 3 population density from 1866 and 1891 to instrument the variable ‘ $\ln dens$ ’. Variable ‘ $\ln MP$ ’ is instrumented by the log of the market potential with population density from 1866; IV3: we use log of car ownership rates from 1999 at the employment area level to instrument the pollution variable,  $NO_X$ . The Endogeneity C-Stat indicates that instrumentation is needed. Besides, the set of instruments is not weak and is exogenous. Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

We also check the statistical difference of the two density coefficients in table 8. We test the null hypothesis that the two coefficients are statistically equal. We obtain a p-value of 0.0365: we thus reject the null hypothesis. Indeed, the two estimated coefficients are statistically different. Moreover, the size of the bias is substantial and represents 13.8% of the estimated coefficient.



Table 8: Comparison between the standard and the extended specification

	IV3	IV4
$\ln dens$	0.0253*** (0.007)	0.0287*** (0.007)
$\ln MP$	0.0373*** (0.009)	0.0589*** (0.007)
$\ln area$	0.0266*** (0.004)	0.0200*** (0.004)
$\ln div$	-0.1881*** (0.047)	-0.2572*** (0.047)
$\ln spe$	0.0285** (0.010)	0.0268* (0.011)
$\ln NO_X$	-0.1031*** (0.021)	-
Industry fixed-effects	Yes	Yes
N	1,485	1,485
Cragg-Donald F-stat	81.291	361.373
Kleibergen-Paap Statistic	51.135	302.372
Hansen J-stat	0.116	0.141
Chi-sq P-value	0.7334	0.7071
Endogeneity C-stat	32.878	29.103
Chi-sq P-value	0.0000	0.0000

Note: IV3: Generalized Method of Moments (GMM) for the extended specification; IV4: Generalized Method of Moments (GMM) for the standard specification; IV3 and IV4: we use log of NUTS 3 population density from 1866 and 1891 to instrument the variable ‘ $\ln dens$ ’. Variable ‘ $\ln MP$ ’ is instrumented by the log of the market potential with population density from 1866; IV3: we use log of car ownership rates from 1999 at the employment area level to instrument the pollution variable,  $NO_X$ . The Endogeneity C-Stat indicates that instrumentation is needed. Besides, the set of instruments is not weak and is exogenous. Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

## 6.2 Estimating the reduction in agglomeration gains: the illustrative case

For the illustrative case, let us assume a representative employment area of 700 square kilometers with a GDP of 5 billion euros and 70,000 workers. We assume the introduction a new structural transportation infrastructure such as a Bus with High Level of Service (BHLS). The infrastructure is expected to create 1,000 new jobs in the employment area.

These hypothesis are totally fictional. The aim of the illustrative case is to provide rough estimates of the reduction in agglomeration economies and to monetarize this loss of wealth.

Due to the implementation of the new transportation infrastructure, the density of the employment area increases by 1.4%. The productivity differential with respect to density is 0.0399% when air pollution is ignored, against 0.0352% when pollution is accounted for. This results in a productivity gain of 28.5 and 25.14 euros per worker, respectively. The agglomeration gains from the 71,000 final workers amount to 2,023,500 euros when pollution is ignored, against 1,784,940 euros when pollution is considered. Therefore, accounting for air pollution reduces the expected agglomeration gains by 13.4%. A 1% increase in  $\text{NO}_x$  emissions reduces the productivity by 0.1%, which corresponds to an economic loss of 238,560 euros for the given level of GDP. The GDP growth expected with the implementation of the new transportation infrastructure is 0.04% when pollution is ignored, against 0.036% when pollution is taken into account. To conclude, considering the aforementioned assumptions, such an infrastructure is expected to generate negligible wealth creation, and an even more negligible one when pollution is accounted for. This illustrative case allows putting into perspective the expected wealth creation resulting from the implementation of a new transportation infrastructure or policy.

## 7 Concluding remarks

This article enlarges the general framework that studies determinants of agglomeration economies by exploring the impact of air pollution on worker productivity. It confirms that pollution has a negative and significant impact on productivity. The results obtained show that taking into account air pollution in agglomeration economies estimations reduces their magnitude by more than 13%. Empirically, the main contribution of this paper is to include a pollution variable in the standard specification of agglomeration economies. The result indicates that air pollution is an omitted variable in standard econometric models estimating agglomeration economies and reduces expected gains. Even if agglomeration economies are substantial when implementing a new transport infrastructure or policy, a part of them should be corrected by the negative environmental

impact from the trips induced by improved accessibility. This paper explicits the general intuition that pollution is harmful to health and that health problems affect negatively labor productivity. It is usually admitted that new transportation infrastructures or policies enhance accessibility and therefore productivity. However, improved accessibility induces new trips which generate increased air pollution. This result provides guidance for policy makers. For this reason, low-carbon transportation infrastructures or policies should be favored to ensure the lowest reduction in the expected agglomeration gains due to air pollution (e.g., car-sharing policies, bike-sharing systems). In addition, policies supporting mobility can be set, such as commuting costs subsidized by firms or mobility learning for young and disadvantaged population.

The results presented are obtained for a specific air pollutant,  $\text{NO}_x$ . Only direct effects are accounted for. It is usually admitted that pollution has cumulative effects on productivity and health. Further work would consist in introducing cumulative effects of air pollution to strengthen our results. In addition, other pollutants can be added to better reproduce air quality and to generalize our findings. Further work could use individual data over several years to control for heterogeneity of workers, in particular with the inclusion of human capital variables such as education. These data would confirm that the results we find are not due to the particular year we use. Moreover, we could investigate the link between the contribution of human capital to agglomeration economies and its variations following the inclusion of pollution.

## References

Briant, A., Combes, P.P. and Lafourcade, M. (2010) Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics* 67, 287–302.

Carson, R. T., Koundouri, P., Nauges. C. (2011) Arsenic Mitigation in Bangladesh: A Household Labor Market Approach, *American Journal of Agricultural Economics* 93, 407–14.

Cervero, R. (2002) Induced travel demand: research design, empirical evidence, and normative policies, *Journal of Planning Literature* 17, 3–20.

Ciccone, A. (2002) Agglomeration effects in Europe, *European Economic Review* 46, 213–227.

Ciccone, A. and Hall, R.E. (1996) Productivity and the density of economic activity, *American Economic Review* 86, 54–70.

CITEPA (2013) *Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France : séries sectorielles et analyses étendues*, Format Secten, Avril.

Combes, P.P., Duranton, G. and Gobillon, L. (2008) Spatial wage disparities: sorting matters, *Journal of Urban Economics* 63, 723–742.

Combes, P.P., Duranton, G., Gobillon, L. and Roux, S. (2010) “Estimating Agglomeration Economies with History, Geology and Worker effects”, in Glaeser, E. (Eds) *Agglomeration Economics*, pp. 15–66. Chicago University Press, Chicago.

Combes, P.P., Lafourcade, M., Thisse, J.F. and Toutain, J.C. (2011) The rise and fall of spatial inequalities in France: A long-run perspective, *Explorations in Economic History* 48, 243–271.

Combes P.P and Lafourcade M. (2012) Revue de la littérature académique quantifiant les effets d’agglomération sur la productivité et l’emploi, Rapport final réalisé pour la Société du Grand Paris.

Combes, P.P., Duranton, G, Gobillon, L., Puga, D. and Roux, S. (2012) The productivity advantages of large cities: Distinguishing agglomeration from firm selection, *Econometrica* 80, 2543–2594.

Currie, J., Hanushek, E., Kahn, M., Neidell, M. and Rivkin, S.G. (2009a) Does pollution increase school absences? *The Review of Economics and Statistics* 91, 682–694.

Currie, J., Neidell, M. and Schmieder, J.F. (2009b) Air Pollution and Infant Health: Lessons from New Jersey, *Journal of Health Economics* 28, 688–703.

DfT (2005) *Transport, wider economic benefits and impacts on GDP*, London, HMSO.

Downs, A. (1992) *Stuck in Traffic: Coping with peak-hour traffic congestion*, The Brookings Institution, Washington DC.

Duranton, G. and Puga, D. (2004) “Micro-foundations of urban agglomeration economies”, in Henderson, J.V. and Thisse, J.F. (Eds) *Handbook of Urban and Regional Economics*, Vol. 4, pp. 2063–2117. North Holland, Amsterdam.

Duranton, G. and Turner, M.A. (2012) Urban growth and transportation, *Review of Economic Studies* 1, 1–36.

Fujita, M., Krugman, P. and Venables, A.J. (1999) *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge/Massachusetts/London: The MIT Press.

Fujita, M. and Thisse, J.F. (2002) *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth*, Cambridge University Press, Cambridge.

Ghosh, R., Joad, J., Benes, I., Dostal, M., Sram, R.J. and Hertz-Picciotto, I. (2012) Ambient nitrogen oxides exposure and early childhood respiratory illnesses, *Environment International* 39, 96–102.

Gibbons, S. and Overman, H. (2009) *Productivity in Transport Evaluation Studies*, Research Report, Department for Transport, London.

Glaeser, E. and Maré, D.C. (2001) Cities and skills, *Journal of Labor Economics* 19, 316–342.

Goodwin, P. (1999) Transformation of transport policy in Great Britain, *Transportation Research A* 33, 655–669.

Graff Zivin, J. and Neidell, M. (2012) The impact of pollution on worker productivity, *American Economic Review* 102, 3652–3673.

Graham, D.J. (2007) Agglomeration, productivity and transport investment, *Journal of Transport Economics and Policy* 41, 317–343.

Graham, D.J., Gibbons, S. and Martin, R. (2009) *Transport Investment and the Distance Decay of Agglomeration Benefits*, Draft Report for the Department of Transport.

Graham, D.J., Melo, P.S., Jiwattanakulpaisarn, P. and Noland, R.B. (2010) Testing for causality between productivity and agglomeration economies, *Journal of Regional Science* 50, 935–951.

Harris, C. (1954) The market as a factor in the localization of industry in the United States, *Annals of the Association of American Geographers* 44, 315–348.

Hanna, R. and Oliva, P. (2011) The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City, *National Bureau of Economic Research Working Paper 17302*.

Head, K. and Mayer, T. (2004) “The Empirics of Agglomeration and Trade”, in Henderson, J.V. and Thisse, J.F. (Eds) *Handbook of Urban and Regional Economics*, Vol. 4, pp. 2118–2171. North Holland, Amsterdam.

Head, K. and Mayer, T. (2006) Regional wage and employment responses to market potential in the EU, *Regional Science and Urban Economics* 36, 573–594.

Henderson, J.V. (1974) The sizes and types of cities, *American Economic Review* 64, 640–656.

Henderson, J.V. (1986) Efficiency of Resource Usage and City Size, *Journal of Urban Economics* 19, 47–70.

- Hills, P. (1996) What is induced traffic? *Transportation* 23, 5–16.
- INSEE (2010) Atlas des zones d'emplois.
- Jacobs, J. (1969) *The Economy of Cities*, Random House, New York.
- Krugman, P.R. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 484–499.
- Latza, U., Gerdes S. and Baur X. (2009) Effects of nitrogen dioxide on human health: Systematic review of experimental and epidemiological studies conducted between 2002 and 2006, *International Journal of Hygiene and Environmental Health* 212, 271–287.
- Lave, L. and Seskin, E. (1970) Air pollution and human health, *Science* 169, 722–733.
- Lavy, V., Ebenstein, A. and Roth, S. (2012) The impact of air pollution on cognitive performance and human capital formation, Working Paper, University of Warwick, 1–38.
- Litman, T. (2011) Generated Traffic and Induced Travel, *Transportation* 71, 38–47.
- Marshall, A. (1890) *Principles of economics*, London, McMillan (8th ed. Published 1920).
- Matas, A., Raymond, J.L. and Roig, J.L. (2013) Wages and Accessibility: The Impact of Transport Infrastructure, *Regional Studies*, (ahead of print), 1–19.
- Melo, P.C., Graham, D.J. and Noland, R.B. (2009) A meta-analysis of estimates of urban agglomeration economies, *Regional Science and Urban Economics* 39, 332–342.
- Moretti, E. (2004) “Human capital externalities in cities”, in Henderson, J.V. and Thisse, J.F. (eds) *Handbook of Urban and Regional Economics*, Vol. 4, pp. 2118–2171. North Holland, Amsterdam.
- Noland, R.B. and Lem, L.L. (2002) A review of the evidence for induced travel and changes in transportation and environmental policy in the US and the UK, *Transportation Research Part D*, 1–26.

Ostro, B.D. (1983) The effects of air pollution on work loss and morbidity, *Journal of environmental economics and management* 10, 371–382

Puga, D. (2010) The magnitude and causes of agglomeration economies, *Journal of Regional Science* 50, 203–219.

Rice, P., Venables, A.J. and Patacchini, E. (2006) Spatial determinants of productivity: analysis for the regions of Great Britain, *Regional Science and Urban Economics* 36, 727–752.

Rosenthal, S.S. and Strange, W.C. (2004) “Evidence on the nature and sources of agglomeration economies”, in Henderson, J.V. and Thisse, J.F. (Eds) *Handbook of Urban and Regional Economics*, Vol. 4, pp. 2118–2171. North Holland, Amsterdam.

Segal, D. (1976) Are there Returns to Scale in City Size?, *Review of Economics and Statistics* 58, 339–350.

Suglia, S., Gryparis, A., Wright, R., Schwartz, J., and R. Wright. (2008) Association of black carbon with cognition among children in a prospective birth cohort study, *American Journal of Epidemiology* 167, 280–286.

Sveikauskas, L. (1975) The productivity of cities, *Quarterly Journal of Economics* 89, 393–413.

Venables, A. (2007) Evaluating urban transport investments. Cost-benefit analysis in the presence of agglomeration and income taxation, *Journal of Transport Economics and Policy* 41, 173–188.

Vickerman, R. (2007) Cost-benefit analysis and large-scale infrastructure projects: state of the art and challenges, *Environment and Planning B* 34, 598–610.

Victoria Department of Transport (2012) *Job density, productivity and the role of transport - An overview of agglomeration benefits from transport investments and implications for the transport portfolio*, State Government Victoria, June.

Wooldridge, J.M. (2009) *Introductory Econometrics: A Modern Approach*, South Western College Publications, fourth ed.



# Appendices

## Appendix A: French employment areas

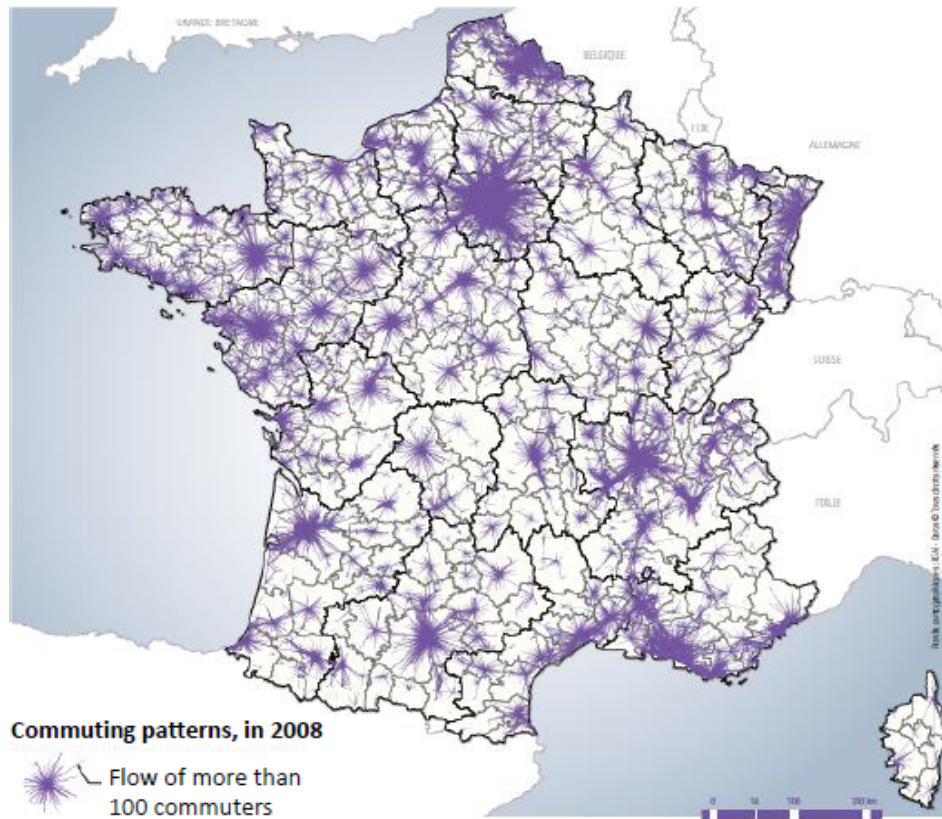


Figure 3: Commuting patterns define French employment areas.  
*Source: INSEE, 2010*

## Appendix B: Interaction coefficient

Only the interaction term is of interest, while the other coefficients are not directly interpretable.

## Appendix C: Correlation matrix of the estimated coefficients

We observe in Table 10 that the correlations between the estimated coefficients are low, which confirms that there is no multicollinearity problem in the extended model presented in Section 5.2

Table 9: Interaction coefficient

	IV5
<i>inter</i>	-0.0186* (0.008)
ln $NO_X$	0.0105 (0.031)
ln <i>dens</i>	-0.0010 (0.025)
ln <i>MP</i>	0.0290*** (0.008)
ln <i>area</i>	0.0330*** (0.004)
ln <i>div</i>	-0.0776 (0.044)
ln <i>spe</i>	0.0375*** (0.010)
Industry fixed-effects	Yes
N	1,485

Note: IV5: Generalized Method of Moments (GMM); we use log of NUTS 3 population density from 1866 and 1891 to instrument the variable ‘ln *dens*’. Variable ‘ln *MP*’ is instrumented by the log of the market potential with population density from 1866; we use log of car ownership rates from 1999 at the employment area level to instrument the pollution variable,  $NO_X$ . Robust standard errors in brackets. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .

Table 10: Correlation matrix of the estimated coefficients

	ln <i>dens</i>	ln <i>MP</i>	ln <i>area</i>	ln <i>div</i>	ln <i>spe</i>	ln $NO_X$
ln <i>dens</i>	1.0000					
ln <i>MP</i>	-0.1152	1.0000				
ln <i>area</i>	0.1695	0.2830	1.0000			
ln <i>div</i>	0.6798	-0.0185	0.1678	1.0000		
ln <i>spe</i>	-0.0795	0.0875	0.0279	-0.1590	1.0000	
ln $NO_X$	0.0959	0.1213	-0.2637	-0.0785	-0.0882	1.0000