

Why do people continue to live near polluted sites? Empirical evidence from Southwestern Europe

Pierre Levasseur

INRA

Katrin Erdlenbruch (✉ katrin.erdenbruch@irstea.fr)

Irstea <https://orcid.org/0000-0002-5470-2636>

Christelle Gramaglia

Irstea

Original Research

Keywords: soil pollution exposure, residential choice, socioeconomic status, environmental inequalities, instrumental variables strategy

Posted Date: March 24th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-169654/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Why do people continue to live near polluted sites?

Empirical evidence from Southwestern Europe

Pierre Levasseur^{1,2}, Katrin Erdlenbruch^{1,3} and Christelle Gramaglia¹

1: G-EAU, Univ Montpellier, INRAE, AgroParisTech, Cirad, IRD, Institut Agro, Montpellier, France

2: UMR SAD-APT, INRAE, AgroParisTech, Univ Paris-Saclay, Paris, France

3: CEE-M, Univ Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France

Corresponding author:

Katrin Erdlenbruch, e-mail: katrin.erdlenbruch@inrae.fr postal address: INRAE 361 rue J-F Breton, 34196 Montpellier cedex 5

Abstract: Poverty is a major determinant for pollution exposure, according to the US location choice literature. In this paper, we assess the impact of socio-economic status on location choices in the European context. Our analysis relies on an original dataset of 1194 households living in polluted and non-polluted areas in three European countries: Spain, Portugal and France. We use instrumental variables strategies to identify the socioeconomic causes of location choices. We show that low education, wealth and income are main reasons for living in polluted areas. We provide several robustness checks testing for the exogeneity of selected instruments. We observe that unobserved heterogeneity tends to understate the impact of socioeconomic status on residence location. Interestingly, we highlight that an important proportion of intermediate social groups (especially young couples) are living in polluted areas, probably because of place attachment and affordable housing facilities. Similarly, we show that middle-income households have lower move-out intentions than other income groups. These latter results contrast the linear vision of environmental inequalities found in the US.

Keywords: *soil pollution exposure; residential choice; socioeconomic status; environmental inequalities, instrumental variables strategy.*

JEL codes: Q53; R23; R58

Author declarations:

- Data and material: The database is anonymous and contains no personal information. All ethical standards concerning data collection and analysis were respected. Anonymized data is available upon request.
- Funding: This study was part of the SoilTakeCare project financed by INTERREG SUDOE funds 2017-2019.
- Competing interests: the authors have no conflicts of interest to declare that are relevant to the content of this article.
- Author's contributions: All authors contributed to the study conception and design. The main data analysis and the first draft of the article was made by the first author. All authors contributed to the writing of the final version of the article.

1. INTRODUCTION

Notwithstanding the strong development of locational choice models in recent decades, empirical evidence for the relationship between socioeconomic status and pollution exposure in Europe remains limited. In the US, the existing literature identifies a negative correlation between socioeconomic status and exposure to pollution (Jerrett et al., 2001, Banzhaf, Ma and Timmins, 2019). Some authors also find evidence of a causal effect (Banzhaf and Walsh 2008): richer households tend to move away from polluted areas while poorer households are more likely to move in; hence pollution is leading to social segregation. In Europe, the literature focused on the role of natural amenities, rather than pollution (but see Chanel et al. 2004), to explain patterns of social segregation. Results are ambiguous: Schaeffer et al. 2016, found that natural amenities increase the mutual segregation between executives and other workers in the French region of Marseille, but not in the region of Grenoble. In the Netherlands, van Duijn and Rouwendal (2013) showed that double earners prefer natural living environments while highly educated households placed more value on historical amenities. De Palma, Picard and Waddell (2007) highlight the important role of the noise disamenity, next to natural amenities and transport amenities, in location choices in the area of Paris. Schaeffer et al. (2016) conclude that location patterns ultimately depend on the interplay between natural and other amenities. According to sociological studies, the dual and linear definition of environmental inequalities should be contrasted in Europe. For instance, in French towns near hazardous industrial wastes, Flanquart, Hellequin and Vallet (2013) observe an over-representation of households with moderate standards of living (and not the poorest). Inequalities resulting from pollution also vary from city to city (Padilla et al. 2014). In the end, there is a non-trivial relationship between environmental and social constraints, explaining what may be termed socio-environmental segregation.¹ In this paper, we study how socio-environmental segregation develops in polluted areas in Southwestern Europe.

¹ Similar terms can be found in the literature: Romero et al. (2012) used the term ‘urban-environmental segregation’ when studying the impact of flood-risk on social segregation. Schaeffer and Tivadar (2018) study ‘segregation based environmental inequalities’ measuring the spatial inequalities of social groups with respect to environmental amenities.

Several theoretical pathways may lead certain social groups to disproportionately live in polluted areas (Banzhaf, Ma and Timmins, 2019). First, industries tend to locate their polluting activities in ‘favorable’ geographic and economic areas, where land is available, the labor force is cheap and transport networks are well developed. Second, households choose their residential location depending on their willingness to pay (Tiebout 1956): even if lower-income households would prefer to live in a cleaner area, they are not willing to pay the higher price. Hence, in the long run, poorer households (as well as certain ethnic groups) end up in more polluted areas (Banzhaf and Walsh, 2013).² Next, households in polluted areas have generally less ability to influence governments and may less easily mobilize against existing rules in favor of industries (Banzhaf, Ma and Timmins, 2019, Gramaglia, 2015). Finally, residents may accept a certain level of pollution exposure in exchange for compensation provided by the polluting firm (Banzhaf, Ma and Timmins 2019, Phillimore and Bell, 2013), such as employment opportunities and direct investments in local amenities (e.g. parks, sports infrastructure and cultural centers).³ Note that the over-representation of households with moderate standards of living observed by Flanquart, Hellequin and Vallet (2013) in French polluted areas may be explained by this pathway. Moreover, place or community attachment might also explain the over-representation of intermediate social classes in polluted areas (De Dominicis et al., 2015).

Given the ambiguities concerning residential preferences in Europe, this article aims to contribute to the literature by quantitatively identifying the factors that determine the probability of living in polluted areas. In addition, we investigate the main determinants of the intention to move out of polluted areas in the next five years. Note that we analyze the *admitted* determinants (i.e. economic and social circumstances), besides the more *hidden* (and often omitted) determinants such as community attachment and risk taking behavior (Dohmen et al., 2011 ; Flanquart, Hellequin and Vallet, 2013 ; De Dominicis et al., 2015). Our methodological approach, based on an instrumental variables (IV) strategy,

² Several multiplier effects can reinforce this pathway (Banzhaf, Ma and Timmins, 2019): people tend to adopt homophile behavior, i.e. they tend to live near other people with similar ethnic, demographic or socioeconomic characteristics (Schirmer, Eggermond and Axhausen, 2014); residents may influence the local amenities of their community, which in turn plays on future arrivals.

³ Compensations may also be set-up by the city hall, which may negotiate subsidies from the industry in exchange for the right to implant a polluting activity.

addresses potential endogeneity issues due to reverse causality and unobserved heterogeneity in the relationship between household socioeconomic status and residential location. To our best knowledge, the use of IV strategy in this context is rare (Bowen, 2002 ; Mitchell and Walker, 2005 ; Xu and Sylwester, 2016), which is another notable contribution of this study. The respondent's height is used as main instrument, arguing that this factor is strongly correlated with socioeconomic status but exogenous to residential location. Our empirical tests validate the conditions that an instrument must satisfy.

Our analysis is based on an original dataset of 1194 households and 2787 individuals in three study areas in Southwestern Europe (Spain, Portugal and France). Comparing the populations from several polluted areas with the populations from similar but non-polluted areas, our results emphasize the presence of strong environmental inequalities. The IV strategy shows that socioeconomic determinants (i.e. education, income and wealth) have strong and negative effects on the probability of living in polluted areas. We clearly observe that observed heterogeneity such as place attachment and local facilities tend to understate the effects of socioeconomic status on residential location. In contrast to the results from the US literature, we confirm sociological assumptions that European polluted areas also seem to be an alternative for young families with lower-middle socioeconomic status who find an interest in affordable housing facilities and employment opportunities, services and infrastructures (see Flanquart, Hellequin and Vallet, 2013). In line with these intuitions, we find that middle-income families seem to see advantages in remaining in polluted areas. As expected, households that are less risk averse and more attached to their community also tend to live in polluted areas. We conclude that such European-specific endogenous non-linearities contribute to make the effect of poverty on risky housing lower.

The structure of the article is the following. In section 2, we describe the database and provide some contextual information about the case studies on which we base our analysis. In section 3, we explain the methods we use to identify the main determinants of the probability of living and continuing to live in polluted areas versus living in cleaner ones. In section 4, we present our results and, in section 5, we conclude and discuss the related public policy implications.

2. DATA

2.1. An original data set

From October 2018 to January 2019, we conducted three household surveys of 1,194 households in France, Portugal and Spain, specifically designed to study the socioeconomic issues of pollution exposure. Using cadastral data, a random selection of housings was employed to guarantee the representativeness of study areas regarding demographic and socioeconomic characteristics. We collected data of 684 households (1589 individuals) in polluted areas and of 510 households (1198 individuals) in corresponding control areas, creating an original comparative dataset, the “Comparative Survey on Pollution Exposure “(CSPE). More precisely, the CSPE is representative of the following polluted areas: Viviez in France (156 households and 293 individuals); the municipality of Estarreja in Portugal (300 households and 739 individuals); and three villages of the Spanish Sierra Minera (Portman, Estrecho de San Ginès and Alumbres) located to the east of Cartagena (228 households and 557 individuals). The non-polluted control areas are: Montbazens in France (138 households and 309 individuals); the municipality of Vagos in Portugal (200 households and 437 individuals); and a group of villages (Portus, Galifa, Perin, La Corona, Cantera and Molinos Marfagones) located to the west of Cartagena in Spain (172 households and 452 individuals). Polluted areas are well-known hotspots of pollution, as confirmed by the literature in geochemistry and mineralogy (Sivry et al. 2010, Patinha et al. 2014, Pérez-Sirvent et al. 2016). Pollution is spread in the polluted municipalities but does not spill over to the control areas. Control areas were selected using region-specific literature. For example, Inácio, Neves and Pereira (2014) and Guihard-Costa et al. (2012) explain that Estarreja and Vagos had the same natural amenities before the installation of the chemical complex in Estarreja. Similarly, the French Institute of Public Health used Montbazens as control area to infer the health effects of pollution exposure in Viviez (Durand, Sauthier and Schwoebel, 2011). Polluted and control areas are sufficiently close so that we are able to suppose that residential choices can be made either in one area or the other. The polluted and control study areas are shown in Figure 1.

It is important to note that this quantitative survey was originally conducted to complete a set of well-documented qualitative interviews. Although the results of the qualitative field campaigns are not directly included in this article, they greatly contributed to our understanding of the study context and issues.

Figure 1: Mapping of polluted and control areas



Source: OSM, authors' computation.

2.2.Context of the study areas

A detailed description of the context of our study areas is given in Appendix A. In a nutshell, our Southwestern European sample makes it possible to observe the three stages that characterize several polluted areas around the world: (i) ex-mining towns (Portman and ESG); (ii) a heavy metal industry undergoing technological reconversion (Viviez); and (iii) active (petro-)chemical complexes (Estarreja, and Alumbres). Table A1 in Appendix A provides means of several important variables of our study in polluted and control areas as well as mean-comparison tests. We also describe at more length the main reasons that motivate residents from Portman/ESG, Viviez, Alumbres and Estarreja to live in the respective areas.

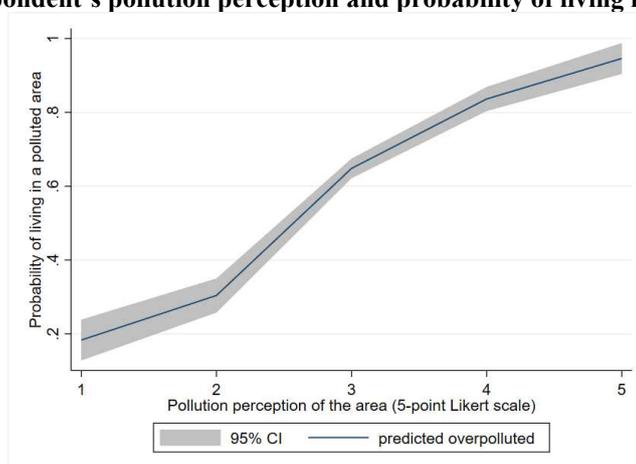
3. METHODS

In this article, we study the determinants of two types of outcome indicators. First, we create a binary variable “living in a polluted area” which takes the value 1 for households living in a study area, and 0 for households living in a control area. This allows us to analyze the reasons why households live in each area. Next, we create a binary response variable that identifies households who plan to move out in the next five years. This allows us to explore potential dynamics in the environmental injustice process. In both cases, we estimate the probability of the outcome indicator being 1, using linear probability models. Referring to the theoretical pathways discussed in Section 2, our empirical analysis especially captures *long-term* effects between household socioeconomic indicators and pollution exposure. Indeed, the industrial sites we study have been settled since at least the 1950s (i.e. for more than two generations). In other words, while the oldest residents might remain for emotional reasons, several current residents already made their (re)location choice as a function of their economic constraints and their willingness to pay for a clean environment.

3.1. Descriptive statistics

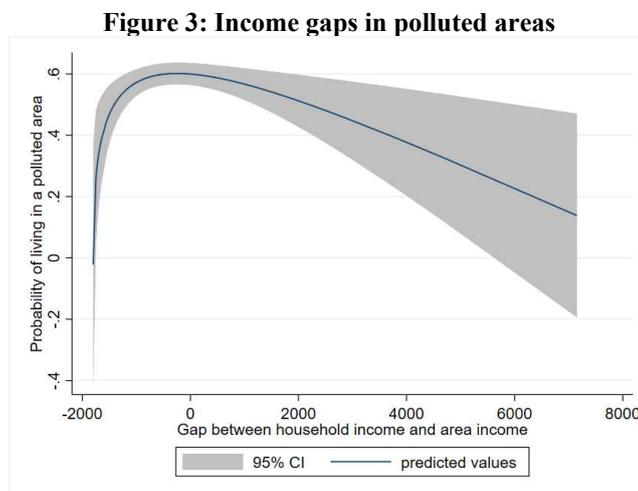
Let us first describe some of the features of our dataset. As shown in Figure 2, the respondent’s pollution perception of the area (measured with a 1-to-5 Likert scale) is strongly correlated with the probability of living in a polluted area. Hence, the question is why people continue to live in polluted areas, even if they know about related pollution issues.

Figure 2: Respondent’s pollution perception and probability of living in polluted areas



Note: Fractional polynomial estimates are used.
Source: Authors’ calculation from the CSPE database.

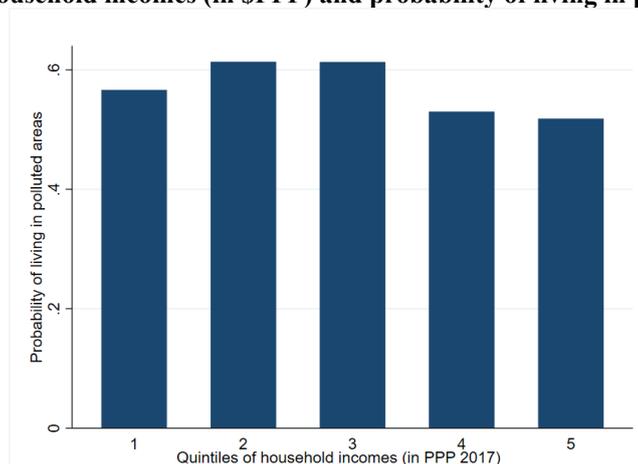
Figure 3 shows that the probability of living in a polluted area decreases when household incomes differ from the average community income. As observed by Guo and Bhat (2007), the lower the absolute gap, the higher the risk of living in a polluted area.



Note: Fractional polynomial estimates are used.
Source: Authors' calculation from the CSPE database.

However, Figure 4 suggests that it is not mainly the poorest households (Q1) who live in polluted areas nor the richest households as expected (Q4 and Q5), but a lower-middle class (Q2 and Q3)⁴. Nonetheless, this result should be interpreted with caution given that there might be endogeneity problems.

Figure 4: Household incomes (in \$PPP) and probability of living in polluted areas



Source: Authors' calculation from the CSPE database.

⁴ We split the sample into five equal groups according to the distribution of household incomes in \$ PPP. Q1 includes the 20% of individuals with the lowest household income, Q2 includes individuals between the 21th percentile and the 40th percentile, Q3 individuals between the 41th and the 60th percentile and Q4 individuals between the 61th and the 80th percentile of household income. Q5 includes the 20% of individuals with the highest household income.

Figures A3 and A4 in Appendix A give some additional information on the way pollution perception is distributed in the sample and how it depends on income and length of residence.

3.2. Empirical challenges

Apart from pollution and industrial features, our sample of polluted and non-polluted areas may initially differ in terms of ecological and historical attractiveness.⁵ In other words, it is impossible to be sure that our selected control areas are perfect counterfactuals of our selected polluted areas. A perfect counterfactual means that these control areas would evolve in the same way as the polluted areas if they had also benefited from the installation of an industrial site (or a mining company). Given that our control areas are potentially imperfect counterfactuals, simply comparing polluted and non-polluted samples might lead to a selection bias. Hence, the main challenge is to deal with endogeneity problems. First, our estimations could be biased because of the presence of reverse causality between household socioeconomic status and pollution exposure. Not only pollution reduces housing prices that potentially attracts poor households, but also polluting industries may emit residuals that are toxic for human health (Landrigan et al., 2018). Daily exposure to these chemical residues may affect the capacity for socioeconomic advancement of residents through loss of productivity. Second, another source of endogeneity may originate from the omission of factors simultaneously correlated with household socioeconomic status and outcome indicators. In our context, we assume that heterogeneous environmental and geographical preferences may bias the estimates. Indeed, these preferences can be simultaneously correlated with socioeconomic status and the unexplained part of (re)location choice, and thus bias estimations. It is widely recognized that different social groups have specific perceptions and preferences about health, pollution and space and thus different ability to pay for desirable community amenities like a clean environment, natural spaces, nice landscapes, quality schools, public safety, employment accessibility and accessible retail outlets (Banzhaf, Ma and Timmins, 2019). To neutralize such a selection bias, we use an IV strategy that allows the effect of household socioeconomic

⁵ Even if we methodically selected the set of control areas, environmental amenities, land availability, soil composition, cultural wealth and other points of interest may differ from those in the polluted areas, even before the arrival of polluting industries.

status to be robustly assessed.⁶ Xu and Sylwester (2016) used a similar approach to assess the impact of area air pollution on emigration flows.

Another challenge is to correct the expected intra-group correlation within polluted villages that could reduce the variance of certain factors, and thus overestimate their significance. Indeed, it is well known that households tend to live among or relocate around groups of households with similar incomes (Schirmer, Eggermond and Axhausen, 2014). Moreover, the presence (or absence) of public facilities is an important predictor of residential choice (Guo and Bhat, 2007), which could reinforce the intra-correlation within an area. To control for the potential intra-group correlations within villages, cluster robust standard errors are systematically estimated at the village level (i.e. the standard errors are not calculated at the individual level but at the village level), see Wooldridge (2003)

3.3. The model

Based on Schirmer, Eggermond and Axhausen (2014), we frame our estimation models on the following dimensions measured at the household level: socioeconomic (income, education and wealth), demographic (age group proportions, gender proportion, marital status) and housing factors (housing size). In exploratory estimations, we also investigate the influence of community-based factors (length of residence and family network) and respondent-based factors (risk taking behavior). Table B1 in the Appendix B describes the explanatory variables that we tested.⁷ More formally, we consider a structural equation with θ designing the functional form of the equation. In the first place, we use a linear

⁶ When endogeneity problems are suspected (e.g. reverse causality and unobserved heterogeneity) and may bias classical estimates such as OLS or Probit, an IV strategy is commonly used in applied economics to perform unbiased fitted coefficients (Wooldridge, 2010). In our study context for instance, the presence of reverse causality and unobserved heterogeneity might overstate or understate the real impact of poverty on the probability of living in a polluted area. By introducing exogenous variations strongly correlated to poverty status but uncorrelated to residential location (i.e. the instrument), IV models only capture the effect of poverty status on pollution exposure that transits through the instrument. If the selected instruments meet both conditions (i.e. being relevant and having no direct link with residential location), one can conclude that fitted coefficients are unbiased and a causal impact is observed. In other words, IV estimates will identify the effect of exogenous variations in poverty status (that transits through the instrument) on the probability of living in polluted areas.

⁷ We also tested other covariates such as housing price and employment indicators, but we do not include them in the model because of collinearity with indicators for socioeconomic status.

probability model (see for example Wooldridge 2010, p. 454). Hence, OLS regressions are performed for Eq.1:

$$\text{Eq.1: } P(y = 1 | \mathbf{x}) = \theta(\mathbf{x}\boldsymbol{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon,$$

with y being the dependent variable ‘living in a polluted area’ and \mathbf{x} the following types of explanatory variables: x_1 socio-economic factors, x_2 demographic factors, x_3 community factors, x_4 individual factors; β_i the corresponding estimated coefficients and ε the error term. Despite the comprehensive set of observed factors included in the analysis, these models potentially remain sensitive to endogeneity problems, mainly due to reverse causality and variations in unobserved individual preferences (landscape preference, geographical location, specific local amenities, place attachment, etc.). Mathematically, β_1 is biased if socioeconomic factors are correlated with ε (i.e. the unexplained part of the variance of the dependent variable). Therefore, to establish a causal inference regarding socioeconomic factors, we apply an IV strategy based on two-stage least square (2SLS) estimations as follows (Wooldridge, 2010):

$$\text{Eq.2: } \begin{cases} P(y = 1 | \mathbf{x}) = \theta(\mathbf{x}\boldsymbol{\beta}) = \beta_0 + \beta_1 \widehat{x}_1 + \beta_2 x_2 + \varepsilon, \\ F(x_1) = \gamma_0 + \gamma_1 z + \gamma_2 x_2 + \rho, \end{cases}$$

with y the dependent variable \mathbf{x} the explanatory variables and ε the error term as above, \widehat{x}_1 the estimated socio-economic factors, F the functional form of the estimation of socio-economic factors, z the instruments, γ_i the corresponding estimated coefficients and ρ the corresponding error term.⁸ As suggested by Angrist and Pischke (2008), we only integrate exogenous control factors in the IV model in order to focus on the causal impacts of household education, wealth and income. Thus, we only control for demographic and housing heterogeneity across households (i.e. age group proportions, gender proportion, marital status, housing size, and country fixed effects). Indeed, the inclusion of potential endogenous control factors (e.g. owning a garden, length of residence, family network and health risk behavior) could bias IV estimates insofar as we cannot be sure whether these factors are determinants or

⁸ Note that β_i and ε are not identical in Eq.1 and Eq.2. This notation is used by convention and hence enhances readability of the equations.

consequences of living in polluted areas. In the first step regression, we linearly regress socioeconomic factors on instruments and covariates. Then, the fitted values of socioeconomic indicators from the first-stage are included in the structural equation to neutralize the unobserved part of the variance of the dependent variable correlated with socioeconomic factors. In other words, assuming exogenous instruments, fitted values of socioeconomic status are suitably independent of omitted factors (ϵ). This means that the model no longer has endogeneity problems and produces consistent estimates.

The case-control design of our database does not allow the use of locality-specific data as instruments such as meteorological variations (e.g. inadequate variation across the study areas). For this reason, we use the reported height (in meters) of the respondent as main IV to instrument socioeconomic factors. The literature abounds with works showing a strong relationship between individual height and socioeconomic status. Indeed, there is a vicious cycle between small height and poverty, namely due to the hazardous consumptions of mothers during pregnancy (e.g. smoking, and alcoholism), micronutrient deficiencies during childhood, schooling and labor market discrimination, and productivity loss due to lower cognitive skills (Banerjee and Duflo, 2011 ; Marmot, 2015). In brief, not only poor households make smaller children but also smaller individuals have lower success in school and employment as well as lower earnings (Magnusson, Rasmussen and Gyllenstein, 2006 ; Case and Paxson, 2008). Even if the negative correlation between poverty and height is expectedly strong, which is the former condition for a valuable instrument, the exogeneity condition is always debatable. Unfortunately, there is no perfect instrument in observational data. Indeed, according to the health economics literature, a child's exposure to pollution might affect growth and then adult height, which may lead to bias IV estimates. For example, Currie and Neidell (2005) observed significant impacts of air pollution peaks on infant mortality risks. In connection with our topic, during the English industrial revolution, Bailey, Hatton and Inwood (2018) found a negative relationship between adult height and intense inhalation of coal smoke in young childhood. However, this study is only correlational and does not identify a causal effect. Epidemiological studies identified a critical window from maternal gestation to early childhood (around age 2) in which a child's physical and intellectual growth is highly dependent on environmental factors

such as feeding practices and mother consumptions (Victora et al., 2008). Moreover, Selevan, Kimmel and Mendola (2000) reported that the health risk related to heavy metals absorption decreases significantly after age 6. Thus, one can assume that a chronic exposure to pollutants occurring up to a certain age may impair body development. Rosales-Rueda and Triyana (2018) achieved in identifying a causal impact of maternal exposure to air pollution peaks on child's height-for-age some years later, but this effect is relatively small and only significant when such risky maternal inhalations occur during the gestation period. Hence, pollution is unlikely to affect the individual growth if the individual starts to be exposed after a certain age. To check for the exogeneity of the respondent's height, we run IV estimates employing several sample restrictions: (i) including only respondents who are not born in the area; (ii) households in which the respondent was at least 4 years old when they moved in the area; (iii) households in which the respondent was at least 18 years old when they moved in the area. To our understanding, if the estimates remain significant in these restricted samples, it means that the effect of pollution exposure on height is mainly due to socioeconomic status. In contrast, if the estimates do not remain significant in the restricted samples, it means that reverse causality problems tend to overstate the effect of poverty on pollution exposure. Furthermore, we provide an additional test of exogeneity by comparing the correlations between the length of area residence and height in polluted areas and control areas. Such a correlation might cast doubt about the identifying assumption.

Next, we consider an additional instrument to implement over-identification tests: the respondent's parental education (a dummy variable identifying if at least one of both parents completed high school). While the correlation between the parent's and children's education is obvious, the exogeneity of parental education with residential location is reasonably questionable. For instance, Bayer, Keohane and Timmins (2009) suggest that one's current location is highly dependent on parental location, the latter being highly correlated to the level of parental education. However, this concern is alleviated when we restrict the sample to individuals who were already 18 years old when they moved in the area. For them, there is no assumed correlation between parental location and current location. Regarding the over-identification test, if we accept the null hypothesis of no correlation between the instruments and

the error term, it means that the latter instrument is valid, based on the assumption that the respondent's height is exogenous, or inversely.

Lastly, to identify the determinants that influence the decision to leave a polluted area, we regress the intention to move out in the next five years on several factors interacted with the fact of living in a polluted area versus living in a cleaner area. We consider that such a model specification (shown in Eq.3) is not particularly affected by endogeneity-related biases. Indeed, an inverse causality bias is not prevalent since the intention to move out is unlikely to affect current income, education or wealth, except if households with move-out intentions work more in order to save money and cover the costs of relocation. While the latter assumption is unlikely to occur, to our extent, we find no reason to think that unobserved factors are simultaneously correlated with socioeconomic status and emigration intentions. In fact, by applying a similar IV strategy as in Eq.2, we do not detect the presence of potential endogeneity problems. Hence, we only run linear probability estimations based on OLS for Eq.3:

$$\text{Eq.3: } P(y = 1 | x) = \theta(\mathbf{x}\boldsymbol{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 [x_1 * x_5] + \varepsilon,$$

where y being the dependent variable 'intention to move out', and \mathbf{x} the following types of explanatory variables: x_1 socio-economic factors, x_2 demographic factors, x_3 community factors, x_4 individual factors, x_5 living in polluted areas; variables x_1 and x_5 are tested in interaction, as shown in the last parenthesis; β_i are the corresponding estimated coefficients and ε the error term.⁹

⁹ Note that β_i , ε and y are not identical in Eq.3 and equations Eq.1 and Eq.2. This notation is used by convention.

4. RESULTS

4.1. Why do people live in polluted areas?

OLS and 2SLS estimates

To test the impact of households' socioeconomic status on the probability of living in polluted areas, we apply an IV strategy based on 2SLS estimations.¹⁰ Table 2 depicts the results for three different indicators of household's socioeconomic status: the number of educated adult members in the household (at least completed a high school diploma), a 7-score wealth index and the logarithm of monthly household income corrected in purchasing power parities (PPP) based on 2017 US dollars. For each socioeconomic indicator, we run linear probability models using OLS, just-identified 2SLS and over-identified 2SLS estimators. In just-identified 2SLS estimates, the respondent's height is used as instrument, whereas in over-identified 2SLS estimates we add the parents' level of education as second instrument.

In Table 2, OLS estimates show very slight correlations between socioeconomic status and residential location. Except for education where the fitted coefficient is significant (Column 1), wealth and income indicators have a negative but non-significant relationship with the probability of living in polluted areas (Columns 4 and 7). By instrumenting socioeconomic status by the respondent's height, just-identified IV estimates show significant and negative impacts of education, wealth and income on the probability of living in polluted settings (Columns 2, 5 and 8). For instance, Column 2 exhibits that one extra-educated adult in the household reduces the probability of living in a polluted area by 46 percentage points (almost 6-times stronger than OLS estimates). Similarly, an increase of monthly incomes by 10% significantly decreases the risk of living in a polluted area by 6.5% (more than 10-times stronger than OLS estimates). Over-identified IV estimates show similar results, even if the instrumented coefficients are a bit lower (Columns 3, 6 and 9).

¹⁰ We also tested the model using IV-Probit estimations. As the results are relatively similar, we only report 2-SLS estimates.

Table 2: Effects of socioeconomic status on the probability of living in polluted areas using OLS, just-identified IV and over-identified IV estimators (linear probability model)

	OLS	Just-identified 2SLS	Over-identified 2SLS	OLS	Just-identified 2SLS	Over-identified 2SLS	OLS	Just-identified 2SLS	Over-identified 2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of children (lower than 17yo)	-0.009 (-0.36)	-0.005 (-0.15)	-0.003 (-0.12)	-0.008 (-0.29)	0.012 (0.49)	0.007 (0.29)	-0.023 (-0.84)	-0.000 (-0.01)	-0.002 (-0.06)
Number of young adults (18-29yo)	0.059* (2.01)	0.276*** (3.62)	0.167** (2.26)	0.014 (0.40)	0.084** (2.29)	0.081** (2.42)	0.012 (0.33)	0.152*** (2.83)	0.111* (1.76)
Number of lower-middle adults (30-44yo)	0.030 (0.94)	0.211*** (2.69)	0.113* (1.86)	-0.005 (-0.15)	0.106* (1.79)	0.089* (1.75)	-0.009 (-0.22)	0.196*** (2.94)	0.129 (1.62)
Number of higher middle adults (45-64yo)	-0.028 (-1.12)	0.099* (1.70)	0.030 (0.66)	-0.048* (-1.86)	0.071 (1.27)	0.047 (1.05)	-0.066 (-1.57)	0.081 (1.51)	0.033 (0.50)
Number of old adults (higher than 65yo)	-0.021 (-0.70)	0.017 (0.42)	-0.001 (-0.04)	-0.031 (-1.06)	-0.060 (-1.39)	-0.048 (-1.40)	-0.058 (-1.53)	0.012 (0.26)	-0.003 (-0.06)
Number of male members	-0.026 (-1.00)	-0.044 (-1.29)	-0.042 (-1.31)	-0.021 (-0.84)	-0.016 (-0.49)	-0.027 (-0.81)	-0.010 (-0.33)	-0.018 (-0.68)	-0.020 (-0.71)
Respondent is in a couple (dummy)	0.016 (0.68)	0.058 (1.57)	0.056* (1.73)	0.022 (0.80)	0.276*** (3.10)	0.257*** (2.58)	0.068* (1.95)	0.241*** (3.88)	0.203*** (2.87)
Respondent is single (dummy)	0.033 (0.66)	0.026 (0.38)	0.040 (0.67)	0.044 (0.93)	0.178*** (2.64)	0.185** (2.49)	0.102* (1.95)	0.182** (2.34)	0.173** (2.54)
Number of rooms	0.042*** (3.20)	0.062*** (3.42)	0.054*** (3.09)	0.043*** (3.49)	0.117*** (4.09)	0.111*** (3.19)	0.043*** (3.11)	0.089*** (4.08)	0.073*** (3.31)
Number of members who obtained a high school grade	-0.088** (-2.74)	-0.464*** (-3.15)	-0.255** (-2.35)						
Wealth index				-0.023 (-1.61)	-0.360*** (-3.16)	-0.291*** (-2.64)			
ln(incomePPP)							-0.056 (-0.96)	-0.646*** (-3.67)	-0.445** (-2.54)
Constant	0.409* (1.86)	0.213 (0.91)	0.288 (1.28)	0.494** (2.25)	1.106*** (4.04)	0.907*** (3.84)	0.815* (1.94)	4.633*** (4.02)	3.309*** (2.97)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,178	1,147	1,098	1,178	1,147	1,098	966	943	911
Hansen J statistic (over-identification test of all instruments) (p-value)			1.692 (0.193)			0.217 (0.641)			1.093 (0.296)
F-statistics on the excluded instruments from first-stage		16.75	16.64		12.77	21.11		20.60	35.03

Notes: (1) In just-identified regressions, only height in meters of the respondent is used as instrument. In over-identified regressions, We add parental education level (at least a high-school degree) of the respondent as second instrument to perform over-identification tests.

(2) Robust standard errors are reported correcting intra-locality correlation. Significance levels are: ***1%, **5%, *10%.

(3) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0. Then, this score is log-transformed by adding 1 for avoiding the generation of missing values (i.e. $\log(0)=.$). We also tested the impact of this variable employing a log-transformation. However, the results remain the same.

Source: Authors' calculation from the CSPE database.

The higher magnitudes of fitted coefficients from IV estimates suggest that OLS estimates strongly understate the impact of socioeconomic status on residential choices, probably due to a selection bias and unobserved heterogeneity. More specifically, the difference between OLS and IV results may be due to unobserved factors correlated with household socioeconomic status. Indeed, according to the sociological literature (e.g. Flanquart, Hellequin and Vallet, 2013), polluted areas have several *omitted* characteristics that may attract lower-middle social classes to live there, such as employment opportunities and accessibility, housing facilities and other advantages (e.g. community satisfaction and attachment). In accordance with this assumption, Table 2 shows that young and lower-middle aged people tend to live more often in polluted areas, either in a couple or as singles. Likewise, coefficients for housing size are significant and positive in all fitted models, corroborating the assumption that households with lower socioeconomic status can live in larger houses at affordable prices in polluted areas. Furthermore, Table B2 in the Appendix, examining the factors that are correlated with the probability of living in a polluted area, clearly confirms the disproportional attraction of young families with lower-middle socioeconomic status. We find a U-inverted relationship between the wealth index or income and the probability of living in polluted areas.¹¹ Moreover, housing size is an important predictor of this probability.¹² Other control factors are tested in these exploratory estimates. We namely observe that community attachment, the presence of a family network and the length of residence in the area increase the probability of living near polluted sites, while risk aversion in the domain of health risks decreases the probability of living there¹³.

¹¹ The turning point is around 1918 \$PPP per month and household, i.e. 1,113 euros in Portugal, 1,230 euros in Spain, and 1,489 euros in France, corresponding to lower-middle incomes in each country.

¹² An alternative model specification adding interaction terms between age groups proportions and housing size shows that housing size particularly increases the probability of living in polluted areas for the 30-45 age group (not shown).

¹³ We measure the environmental health risk aversion by asking the respondent to evaluate on a 1-to-5 Likert scale his/her willingness to live in a polluted area that may decrease by 5 years old his/her life expectancy.

Validation of the IV strategy

Empirical tests suggest that our instruments satisfy the two requirements of suitable instruments, particularly for the respondent's height (Wooldridge 2010).¹⁴ First, both instruments are non-weak predictors of the endogenous variable, conditional on control variables. As shown in Table B3 in the Appendix, even after controlling for exogenous covariates, the height of the respondent and the education of his/her parents are both significant predictors of household socioeconomic status. It is reassuring to see in Table 2 that all first-stage F-statistics on the excluded instrument are relatively high. Second, the exclusion restriction assumption can be partially tested when the endogenous regressor is over-identified. Over-identification tests suggest the absence of correlation between the error terms of the structural equation and the instruments (Table 2). Likewise, Figure B1 in the Appendix clearly shows that a longer exposure to soil pollution does not affect height differently compared to control areas. In other words, residents from polluted areas are not smaller because of a longer exposure to pollution but because of other reasons such as poverty. In addition, we check the robustness of IV estimates by restricting the sample to: (i) households for which the respondent is not born in the area; (ii) households for which the respondent was at least 4 years old when he/she moved in the area; (iii) households for which the respondent was at least 18 years old when he/she moved in the area. Table 3 exhibits these IV estimates based on restricted samples. Interestingly, Table 3 reports lower effects than Table 2, suggesting that a potential reverse causality could slightly overstate the impact of poverty on pollution exposure. However, by neutralizing all suspected links between the instruments and pollution exposure, the negative impact of socioeconomic status on the probability of living in polluted areas remains strong and significant, even in just-identified 2SLS estimates where only the respondent's height is used as instrument. We observe in these bias-robust estimates that one extra-educated adult member decreases

¹⁴ To be suitable, an instrument must meet two conditions: (i) it must be a non-weak predictor of the endogenous variable conditional on control variables; and (ii) it must not be directly related to the error component in the structural equation (i.e. not be correlated with the unexplained part of the probability of living in a polluted area). The second condition, called exclusion restriction assumption, means that our instruments should not directly correlate with the probability of living in a polluted area through channels other than the household socioeconomic status (Wooldridge, 2010).

up to 40 percentage points the risk of living in polluted areas. Similarly, one extra owned asset and 60% additional income decrease this risk by around 30 percentage points.

Table 3: Robustness checks for the IV exogeneity condition (linear probability model)

	No born in the area		At least 4 yo when moved in the area		At least 18 yo when moved in the area	
	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>
Number of members who obtained a high school grade	-0.401*** (-3.21)	-0.353*** (-4.44)	-0.363*** (-2.94)	-0.340*** (-3.96)	-0.254** (-2.09)	-0.274*** (-3.66)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	664	628	622	586	531	502
Hansen J statistic (p-value)		0.009 (0.926)		0.046 (0.830)		0.346 (0.557)
F-statistics on the excluded instruments	18.04	12.48	19.07	12.28	14.52	11.86
	No born in the area		At least 4 yo when moved in the area		At least 18 yo when moved in the area	
	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>
Wealth index	-0.321*** (-3.22)	-0.366*** (-4.15)	-0.269*** (-3.62)	-0.321*** (-4.18)	-0.174** (-2.25)	-0.225*** (-2.91)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	664	628	622	586	531	502
Hansen J statistic (p-value)		0.752 (0.386)		1.557 (0.212)		1.798 (0.180)
F-statistics on the excluded instruments	11.38	20.56	14.13	18.27	11.46	15.55
	No born in the area		At least 4 yo when moved in the area		At least 18 yo when moved in the area	
	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>	<i>Just-identified IV</i>	<i>Over-identified IV</i>
ln(incomePPP)	-0.480*** (-3.98)	-0.523*** (-3.84)	-0.476*** (-3.48)	-0.549*** (-3.68)	-0.303* (-1.79)	-0.437*** (-2.76)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	537	514	499	476	418	401
Hansen J statistic (p-value)		0.640 (0.424)		1.113 (0.292)		2.544 (0.111)
F-statistics on the excluded instruments	34.49	28.24	29.44	23.98	24.81	22.70

Notes: (1) In just-identified regressions, only height in meters of the respondent is used as instrument. In over-identified regressions, We add parental education level (at least a high-school degree) of the respondent as second instrument to perform over-identification tests.
(2) Robust standard errors are reported correcting intra-locality correlation. Significance levels are: ***1%, **5%, *10%,
(3) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0. Then, this score is log-transformed by adding 1 for avoiding the generation of missing values (i.e. $\log(0)=.$). We also tested the impact of this variable employing a log-transformation. However, the results remain the same.

Source: Authors' calculation from the CSPE database.

Furthermore, by running a reduced-form regression of the dependent variable on the instruments and covariates (Table B4 of the Appendix), we are able to appreciate the expression of the causal effect of interest. Indeed, we find a significant correlation between our instruments and the probability of living

in a polluted area, which is proportional to the effect of household socioeconomic status on residential location. This means that when household socioeconomic status is omitted, we continue to observe its influence on the dependent variable through the instruments.¹⁵ Consequently, we can assume that our instruments are not only strong (i.e. meet the first requirement), but also exogenous (i.e. meet the second requirement).

4.2. What factors motivate people to move out of polluted areas?

Table 4 reports OLS estimates regarding the main identified determinants of the intention to move-out in the next 5 years (the full table is available in Table B5 of the Appendix).¹⁶ Several divergences appear between polluted areas and control areas. First, compared to control areas, living in polluted areas increases the intention to move out in the next five years by 3.4 percentage points (Column 1). Focusing on interaction terms, we observe that both the wealth index and income groups affect the intention to leave a polluted area. Column 2 of Table 4 exhibits that one extra-owned asset increases the move-out intention by 2.2 percentage points in polluted areas. In contrast, Table B5 in the Appendix shows that such an extra asset decreases this intention by 1.9 percentage points in control areas (i.e. a gap equal to 4.1 percentage points between both areas). Column 2 of Table 4 also hypothetically indicates that, among extremely deprived households (with not even one owned asset), living in a polluted area does not affect move-out intentions. In Column 3, we find that, in polluted areas, households belonging to the third (Q3) and the fourth quintiles (Q4) have lower move-out intentions compared to households belonging to the richest income category (Q5), by around -16 and -14 percentage points respectively. In control areas though, income is not correlated to this intention (Table B5 in the Appendix). These results emphasize

¹⁵ As Angrist and Pischke (2008, p.157) argue, “if you can’t see the causal relation of interest in the reduced form, it’s probably not there”.

¹⁶ Other control variables available in Table B5 of the Appendix do not influence move-out intentions differently between polluted and non-polluted areas. For instance, the attractiveness of an area and the perception of social cohesion are two of such factors, probably because of a collinearity problem with residential location. In the same way, household education, marital status and the presence of a family network affect the intention to move out independently of the residential location. Anywhere, living close his/her family reduces move-out intentions, whereas being single and educated increases mobility intention, probably due to better social and professional opportunities elsewhere.

a nonlinear relationship between household socioeconomic status and the intention to leave a polluted area.

Table 4: Main factors correlated to the intention to move out in the next 5 years (OLS estimates)

	(1)	(2)	(3)	(4)	(5)
Polluted area (dummy)	0.034*	-0.045	0.139**	0.292***	0.229**
	(1.90)	(-1.09)	(2.13)	(3.18)	(2.47)
Wealth*PollutedArea		0.022*			
		(1.93)			
IncomeQ1*PollutedArea			-0.105	-0.058	-0.039
			(-1.67)	(-1.03)	(-0.68)
IncomeQ2*PollutedArea			-0.078	-0.053	-0.044
			(-1.18)	(-0.85)	(-0.70)
IncomeQ3*PollutedArea			-0.166*	-0.165**	-0.166**
			(-2.02)	(-2.09)	(-2.12)
IncomeQ4*PollutedArea			-0.136**	-0.130**	-0.124**
			(-2.40)	(-2.44)	(-2.43)
AverageAge*PollutedArea				-0.003***	-0.003***
				(-3.55)	(-3.50)
Association*PollutedArea					0.073*
					(1.65)
CommunityParties*PollutedArea					0.113**
					(2.29)
Constant	0.514***	0.585***	0.499***	0.411***	0.449***
	(5.37)	(5.31)	(5.11)	(4.16)	(4.70)
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	1,099	1,099	900	900	900
R-square	0.110	0.114	0.133	0.139	0.150

Notes: (1) Standard errors are robust to intra-group correlation. Significance levels are: ***1%, **5%, *10%,

(2) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0.

Source: Authors' calculation from the CSPE database.

Regarding demographic factors, Column 4 in Table 4 shows that the average age of adult household members significantly reduces the motivation to move out in households living in polluted areas (10 extra years in age reduces move-out intentions by 3 percentage points), which is consistent with the literature (Lee and Waddell, 2010). Community attachment might also explain why old people decide to continue living in the same place. Finally, we find a surprising result concerning community involvement. While participation in community events increases the intention to stay in a cleaner control areas by 5.7 percentage points (Table B5 in the Appendix), the same participation increases the intention to leave a polluted area by 11.3 percentage points. These findings highlight the presence of a link between community involvement and sensitivity to pollution, as mentioned by Chanel et al. (2004).

Moreover, these results might emphasize the level of social exclusion of several households who prefer to stay than to move out.¹⁷

5. DISCUSSION AND CONCLUSION

Our study identified the main determinants that explain why households live and continue to live near polluted sites in Southwestern Europe. We implemented an original IV procedure based on the respondent's height, which is rarely done in the connected literature. Globally, our results corroborate the existence of socio-environmental inequalities in the European context. First, we find that household education strongly reduces the risk of living in polluted areas. In line with the results from the US literature, environmental disamenities tend to ward off educated households, whereas environmental amenities attract these population groups (Waltert and Schläpfer, 2010). Negative impacts of wealth and income are also highlighted, suggesting that polluted areas also represent an environment-poverty trap in European countries (Ikefuji and Horii, 2007). Interestingly, contrary to the US, European specificities make ambiguous the correlation between socioeconomic status and pollution exposure. Indeed, our data distribution suggests that the social segregation is not as linear as in the US. Additional exploratory analyses show that a lower-middle income class disproportionately lives in polluted areas in Southwestern Europe. Note that housing size and place attachment significantly increases the probability of living in polluted areas for this social group. In addition, there are higher proportions of couples and young adults who live in polluted areas to benefit from bigger houses. Hence, as pertinently discussed by Flanquart, Hellequin and Vallet (2013), polluted areas may constitute an acceptable residential alternative for households with moderate standards of living insofar as such towns provide several amenities at an affordable price. The overrepresentation of a lower-middle class in larger houses also matches the assumption of Banzhaf, Ma and Timmins (2019) about the existence of compensations and benefits for households who accept to live near a polluting industry. We conclude that such unobserved local

¹⁷ One could also assume that this latter result reflects the presence of reverse causality, if individuals with move-out intentions become less engaged in community activities as they plan to cut ties soon. However, the fact that such a behavior is not observed in clean areas clearly invalidates this possibility.

amenities might be a source of endogeneity and are likely to contribute to understate the effect of socioeconomic status on residential location when classic estimators are used.

In terms of move-out intentions, we detect some nonlinearities that might modify the linear vision of environmental inequalities. In accordance with the mainstream theory, it is particularly the richest households who plan to move out, probably because of their greater financial capacities and employment opportunities. Inversely, the most materially deprived households have no move-out intentions given their limited funding capacities and opportunities. Moreover, among polluted areas, we observe that people who are not involved in the community life have lower move-out intentions than people who are involved in community life. To us, these findings suggest the disproportionate presence of socially excluded groups in polluted areas with no move-out intentions. Our results emphasize the fact that the general equilibrium theorized by Tiebout (1956) has still not been reached in our sample of polluted areas. However, between the two socioeconomic antipodes, there is a lower middle class with strong intentions to remain in polluted areas. As we previously suspected, one can assume that polluted areas in Southwestern Europe provide some amenities that particularly attract young families with moderate standards. Another interesting result underlines the importance of aging as significant determinant of the intention to remain in polluted areas. This result is in line with the sociological literature. In the US context, Shriver and Kennedy (2005, p.495) argue that “long-term residents express less concern over environmental hazards because they are far more attached to their local communities”. In France, Flanquart, Hellequin and Vallet (2013) observe that populations living alongside hazardous industrial sites feel a strong community attachment.

The main limitation of this study is linked to the fact that it is measuring local effects. Indeed, strictly speaking, our results are only valid for our study areas in Southwestern Europe. However, the industrial heterogeneity of our sample makes our estimate generalizable to a wide spectrum of polluted areas. Of course, further analyses of the determinants of the socio-environmental segregation in Europe should be conducted in other contexts (e.g. different case studies and different outcome indicators).

The fact that hazardous polluted areas tend to become economically attractive for lower socioeconomic groups may be a dramatic public health issue, given that lower social classes tend to have more children. Additional studies should also assess the health and productivity effects of pollution exposure in addition to the influence of socioeconomic status on such effects. Finally, our results imply that health policies or recommendations for averting behaviour should be targeted in particular towards lower-middle class families, which are the most likely to find an interest in employment and housing opportunities in polluted areas.

REFERENCES

- ANGRIST Joshua D., PISCHKE Jörn-Steffen, 2008, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 393 p.
- BAILEY Roy E., HATTON Timothy J., INWOOD Kris, 2018, "Atmospheric Pollution, Health, and Height in Late Nineteenth Century Britain", *The Journal of Economic History*, 78(4), pp. 1210–1247.
- BANERJEE Abhijit, DUFLO Esther, 2011, *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, First Edition edition, New York, PublicAffairs, 320 p.
- BANOS-GONZALES Isabel, BAÑOS PAEZ Pedro, 2013, *Portmán: De el Portus Magnus del Mediterráneo Occidental a la Bahía Aterrada | Ediciones de la Universidad de Murcia (Editum)*, Ediciones de la Universidad de Murcia, Murcia, España, Baños Paez.
- BANZHAF H. Spencer, WALSH Randall P., 2008, "Do people vote with their feet? An empirical test of Tiebout", *American Economic Review*, 98(3), pp. 843–863.
- BANZHAF H. Spencer, WALSH Randall P., 2013, "Segregation and Tiebout sorting: The link between place-based investments and neighborhood tipping", *Journal of Urban Economics*, 74, pp. 83–98.
- BANZHAF Spencer, MA Lala, TIMMINS Christopher, 2019, "Environmental Justice: the Economics of Race, Place, and Pollution", *The Journal of Economic Perspectives: A Journal of the American Economic Association*, 33(1), pp. 185–208.
- BAYER Patrick, KEOHANE Nathaniel, TIMMINS Christopher, 2009, "Migration and hedonic valuation: The case of air quality", *Journal of Environmental Economics and Management*, 58(1), pp. 1–14.
- BOWEN William, 2002, "An analytical review of environmental justice research: what do we really know?", *Environmental Management*, 29(1), pp. 3–15.
- CASE Anne, PAXSON Christina, 2008, "Stature and status: Height, ability, and labor market outcomes", *The journal of political economy*, 116(3), pp. 499–532.
- CHANEL Olivier, FAUGERE Elsa, GENIAUX Ghislain, KAST Robert, LUCHINI Stéphane, SCAPECCHI Pascale, 2004, "Valorisation économique des effets de la pollution atmosphérique: Résultats d'une enquête contextuelle", *Revue économique*, 55(1), pp. 65–92.
- CONESA Héctor M., SCHULIN Rainer, NOWACK Bernd, 2008, "Mining landscape: A cultural tourist opportunity or an environmental problem?: The study case of the Cartagena–La Unión Mining District (SE Spain)", *Ecological Economics*, 64(4), pp. 690–700.
- CURRIE Janet, NEIDELL Matthew, 2005, "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?", *The Quarterly Journal of Economics*, 120(3), pp. 1003–1030.
- DE DOMINICIS Stefano, FORNARA Ferdinando, GANUCCI CANCELLIERI Uberta, TWIGGER-ROSS Clare, BONAIUTO Marino, 2015, "We are at risk, and so what? Place attachment, environmental risk perceptions and preventive coping behaviours", *Journal of Environmental Psychology*, 43, pp. 66–78.
- DE PALMA André, PICARD Nathalie, WADDELL Paul, 2007, "Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region", *Journal of Urban Economics*, 62(2), pp. 204–230.
- DOHMEN Thomas, HUFFMAN David, SCHUPP Jürgen, FALK Armin, SUNDE Uwe, WAGNER Gert G., 2011, "Individual risk attitudes: measurement, determinants, and behavioral consequences", *Journal of the European Economic Association*, 9(3), pp. 522–550.
- DURAND Cécile, SAUTHIER Nicolas, SCHWOEBEL Valérie, 2011, "Évaluation de l'exposition à des sols pollués au plomb, au cadmium et à l'arsenic en Aveyron", *Rapport de l'Institut de Veille Sanitaire (InVS)*, France, CIRE/Production scientifique InVS.
- FLANQUART Hervé, HELLEQUIN Anne-Peggy, VALLET Pascal, 2013, "Living alongside hazardous factories: risk, choice and necessity", *Health, Risk & Society*, 15(8), pp. 663–680.
- GUIHARD-COSTA Anne-Marie, INACIO Manuela, VALENTE Sandra, FERREIRA DA SILVA Eduardo, 2012, "Méthode d'étude spatialisée des effets de la contamination industrielle sur la santé des populations locales, région d'Estarreja (Portugal)", *Sud-Ouest européen. Revue géographique des Pyrénées et du Sud-Ouest*, 33, pp. 69–76.
- GUO Jessica Y., BHAT Chandra R., 2007, "Operationalizing the concept of neighborhood: Application to residential location choice analysis", *Journal of Transport Geography*, 15(1), pp. 31–45.
- IKEFUJI Masako, HORII Ryo, 2007, "Wealth Heterogeneity and Escape from the Poverty–Environment Trap", *Journal of Public Economic Theory*, 9(6), pp. 1041–1068.

- INÁCIO M, NEVES O, PEREIRA V, 2014, “Acumulação de metais pesados em forragens e produtos agrícolas em Podzóis de um local industrial Português”, *Comunicações Geológicas*, 101, pp. 1019–1022.
- JERRETT Michael, BURNETT Richard T, KANAROGLOU Pavlos, EYLES John, FINKELSTEIN Norm, GIOVIS Chris, BROOK Jeffrey R, 2001, “A GIS–Environmental Justice Analysis of Particulate Air Pollution in Hamilton, Canada”, *Environment and Planning A: Economy and Space*, 33(6), pp. 955–973.
- LANDRIGAN Philip J., FULLER Richard, ACOSTA Nereus J. R., ADEYI Olusoji, ARNOLD Robert, BASU Niladri (Nil), BALDÉ Abdoulaye Bibi, BERTOLLINI Roberto, BOSE-O'REILLY Stephan, BOUFFORD Jo Ivey, BREYSSE Patrick N., CHILES Thomas, MAHIDOL Chulabhorn, COLL-SECK Awa M., CROPPER Maureen L., FOBIL Julius, FUSTER Valentin, GREENSTONE Michael, HAINES Andy, HANRAHAN David, HUNTER David, KHARE Mukesh, KRUPNICK Alan, LANPHEAR Bruce, LOHANI Bindu, MARTIN Keith, MATHIASSEN Karen V., MCTEER Maureen A., MURRAY Christopher J. L., NDAHIMANANJARA Johanita D., PERERA Frederica, POTOČNIK Janez, PREKER Alexander S., RAMESH Jairam, ROCKSTRÖM Johan, SALINAS Carlos, SAMSON Leona D., SANDILYA Karti, SLY Peter D., SMITH Kirk R., STEINER Achim, STEWART Richard B., SUK William A., SCHAYCK Onno C. P. van, YADAMA Gautam N., YUMKELLA Kandeh, ZHONG Ma, 2018, “The Lancet Commission on pollution and health”, *The Lancet*, 391(10119), pp. 462–512.
- LEE Brian H. Y., WADDELL Paul, 2010, “Residential mobility and location choice: a nested logit model with sampling of alternatives”, *Transportation*, 37(4), pp. 587–601.
- MAGNUSSON Patrik K. E., RASMUSSEN Finn, GYLLENSTEN Ulf B., 2006, “Height at age 18 years is a strong predictor of attained education later in life: cohort study of over 950,000 Swedish men”, *International Journal of Epidemiology*, 35(3), pp. 658–663.
- MARMOT Michael, 2015, “The health gap: the challenge of an unequal world”, *The Lancet*, 386(10011), pp. 2442–2444.
- MITCHELL Gordon, WALKER Gordon, 2005, “Methodological issues in the assessment of environmental equity and environmental justice”, in *Sustainable Urban Development: The environmental assessment methods*, Taylor & Francis, pp. 447–472.
- PADILLA C.M., KIHAL-TALANTIKITE W., VIEIRA V.M., ROSSELLO P., LE NIR G., ZMIRO-NAVIER D., DEGUEN S. 2014, “Air quality and social deprivation in four French metropolitan areas—A localized spatio-temporal environmental inequality analysis”, *Environmental Research*, 134, pp. 315–324.
- PATINHA C., REIS A. P., DIAS A.C., ABDULJELIL A.A., NOACK Y., ROBERT S., CAVE M., FERREIRA DA SILVA E., 2014, “The mobility and human oral bioaccessibility of Zn and Pb in urban dusts of Estarreja (N Portugal)”, *Environmental Geochemistry and Health*, 37, pp. 115–131.
- PÉREZ-SIRVENT C., HERNÁNDEZ-PÉREZ C., MARTÍNEZ-SÁNCHEZ M.J., GARCÍA-LORENZO M., BECH J., 2016, “Geochemical characterisation of surface waters, topsoils and efflorescences in a historic metal-mining area in Spain”, *Journal of soils and sediments*, 16, pp. 1238–1252.
- PHILLIMORE Peter, BELL Patricia, 2013, “Manufacturing loss: Nostalgia and risk in Ludwigshafen”, *Focaal*, 2013(67), pp. 107–120.
- ROMERO Y., VÁSQUEZ, A., FUENTES C., SALGADO M., SCHMIDT A., BANZHAF E., 2012, “Assessing urban environmental segregation (UES). The case of Santiago de Chile”, *Ecological Indicators*, 23, pp. 76–87.
- ROSALES-RUEDA Maria, TRIYANA Margaret, 2018, “The Persistent Effects of Early-Life Exposure to Air Pollution: Evidence from the Indonesian Forest Fires”, *Journal of Human Resources*, pp. 0117-8497R1.
- SCHAEFFER Y., CREMER-SCHULTE D., TARTIU C., TIVADAR M., 2016, “Natural amenity-driven segregation: Evidence from location choices in French metropolitan areas”, *Ecological Economics*, 130, pp. 37–52.
- SCHAEFFER Y., TIVADAR M., 2018, “Measuring Environmental Inequalities: Insights from the Residential Segregation Literature”, *Ecological Economics*, 164, <https://doi.org/10.1016/j.ecolecon.2019.05.009>
- SCHIRMER Patrick M., EGGERMOND Michael A. B. van, AXHAUSEN Kay W., 2014, “The role of location in residential location choice models: a review of literature”, *Journal of Transport and Land Use*, 7(2), pp. 3–21.
- SELEVAN, KIMMEL, MENDOLA, 2000, “Identifying critical windows of exposure for children’s health.”, *Environmental Health Perspectives*, 108 (suppl 3), pp. 451–455.
- SIVRY, Y., MUNOZ, M., SAPIN-DIDIER, V., RIOTTE, J., DENAIX, L. DE PARSEVAL, P., DESTRIGNEVILLE, C., DUPRÉ, B. 2010, “Multimetallic contamination from Zn-ore smelter: solid speciation and potential mobility in riverine floodbank soils of the upper Lot River (SW France)”, *European Journal of Mineralogy*, 22(5), pp. 679–691.
- SHRIVER Thomas E., KENNEDY Dennis K., 2005, “Contested Environmental Hazards and Community Conflict Over Relocation*”, *Rural Sociology*, 70(4), pp. 491–513.
- TIEBOUT Charles M., 1956, “A Pure Theory of Local Expenditures”, *Journal of Political Economy*, 64(5), pp. 416–424.

VAN DUJN Mark, **ROUWENDAL** Jan, 2013, “Cultural heritage and the location choice of Dutch households in a residential sorting model”, *Journal of Economic Geography*, 13(3), pp. 473–500.

VICTORA Cesar G, **ADAIR** Linda, **FALL** Caroline, **HALLAL** Pedro C, **MARTORELL** Reynaldo, **RICHTER** Linda, **SACHDEV** Harshpal Singh, 2008, “Maternal and child undernutrition: consequences for adult health and human capital”, *The Lancet*, 371(9609), pp. 340–357.

WALTERT Fabian, **SCHLÄPFER** Felix, 2010, “Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies”, *Ecological Economics*, 70(2), pp. 141–152.

WOOLDRIDGE Jeffrey M., 2003, “Cluster-Sample Methods in Applied Econometrics”, *American Economic Review*, 93(2), pp. 133–138.

WOOLDRIDGE Jeffrey M., 2010, *Econometric Analysis of Cross Section and Panel Data*, Second Edition, Cambridge MA, The MIT Press.

XU Xu, **SYLWESTER** Kevin, 2016, “Environmental Quality and International Migration”, *Kyklos*, 69(1), pp. 157–180.

APPENDIX

APPENDIX A: The context of the study areas

The three study areas have different mining and industrial histories that make their comparison generalizable to a wide spectrum of pollution contexts. In the following, we describe some key characteristics of each area.

The Spanish Sierra Minera is an ex-mining site that was particularly active between 1957 and 1990 due to the activity of a multinational company. Soils in the area show high concentrations of zinc, lead and cadmium. Since the decline of mining, few industrial alternatives have been set up and the development of tourism remains uncertain (Conesa, Schulín and Nowack, 2008 ; Banos-Gonzales and Baños Paez, 2013). By comparing Portman and Estrecho de San Ginès ESG to other towns located to the west of Cartagena (control group), Table 1 shows that lower housing prices, household incomes and employment rates characterize such ex-mining sites. Table 1 also identifies the lower perceived availability of services and retail outlets in these areas.

Alumbres is a small town located at the foot of the Sierra Minera (between Cartagena and La Union). This small town has prospered alongside the gradual development of a large petrochemical complex since 1950. Today, this industrial site includes an oil refinery, a gas plant, an electric power station that transforms fuel oils and gas, a factory producing white minerals oils, natural sulfonates and sulfuric acid, a fertilizer industry, and, a producer of lubrication bases. Alumbres is exposed to toxic winds of heavy metal residuals. As shown in Table 1, there is no significant difference between Alumbres and Molinos Marfagones (control group) in terms of the price of housing, employment, unemployment and perceived availability of retail outlets.

The Portuguese region of Estarreja has hosted an active industrial site since 1946. First, ammoniac, chlorine-sodium and PVC manufactures settled in Estarreja in 1946, 1956 and 1960, respectively. Then, since 1977, several petrochemical industries have begun their activity. Today, Estarreja hosts six complementary industries producing a large number of chemical products and other derived goods.

Water canals and ditches around the factory transport heavy metals and organic compounds. For instance, high concentrations of lead, mercury, arsenic and benzene have been found in the area. On the other hand, the presence of the industrial complex has made the area more dynamic and has improved the average socioeconomic and demographic characteristics of the area (Inácio, Neves and Pereira, 2014 ; Guihard-Costa et al., 2012). As suggested by Table 1, the municipality of Estarreja has better average characteristics than the municipality of Vagos (control area). Compared to the control area, residents are significantly richer in Estarreja, besides having better employment indicators and public facilities.

The case of Viviez in France marks the transition between a zinc smelting and a modern industry based on the processing of zinc and the recycling of industrial wastes. In 1855, a zinc smelter settled in Viviez because of its proximity to coal mines and rail facilities. In 1871, an international company undertook large-scale industrialization of the site by extracting, transforming and exporting zinc. In 1922, the site became a pioneer in adopting electrolytic techniques to chemically extract zinc from zinc blende. Although zinc extraction ended in 1987, the company continues to process zinc. In addition, the company helped develop new industries, namely the recycling of cadmium residuals and plastics. As a result of the zinc melting activity, the soils in the areas are contaminated with high concentrations of lead, cadmium and arsenic (Durand, Sauthier and Schwoebel, 2011). Considering the economic and community indicators listed in Table 1, Viviez looks more like Portman/ESG (an ex-mining site) than Alumbres and Estarreja (active industrial sites). Compared to Montbazens (control area), Viviez has lower housing prices, household incomes and employment rates, as well as a lower perceived availability of public services and retail outlets.

Based on the CSPE database, Table A1 provides mean-comparison tests between polluted sites and their respective control areas.

Table A1: Mean-comparison tests between polluted and control areas

	Portman/ESG vs. control	Viviez vs. control	Alumbres vs. control	Estarreja vs. control
<i>Economic characteristics</i>				
Declared housing price (euros)	-37326***	-81800***	-4187	-1089
Monthly renting price (euros)	-62***	-104***	-26	6.2
Monthly total income (euros)	-504***	-561***	-139	268***
Employment (proportion)	-0.20***	-0.06***	-0.07	0.16***
Unemployment and inactivity (proportion)	0.20***	0.14**	0.14	-0.17***
<i>Local advantages</i>				
Perception of area attractiveness (1-to-5 scale)	-0.16	-1.21***	-0.53***	0.02
Perception of public services availability (1-to-5 scale)	-1.06***	-0.75***	-0.41**	0.20**
Perception of shops and retails availability (1-to-5 scale)	-0.36*	-1.09***	-0.25	0.26***
<i>Community life and involvement</i>				
Perception of social cohesion (1-to-5 scale)	-0.2	-0.51***	-0.40**	-0.79***
Regularly participating in community events (proportion)	0.06	-0.12**	0.20**	0.05
Involved in a local association (proportion)	-0.11*	-0.09*	-0.04	0.07**

Notes: (1) Mean comparison tests ($H_0: \text{diff}=0$) are corrected for unequal variances when the standard deviation (SD) ratio is significantly different to 1 ($H_0: \text{ratio}=1$). Bilateral significance levels: ***1%, **5%, *10%,
(2) Monthly renting price is hypothetical for housing owners.

Source: Authors' calculation from the CSPE database.

Additional descriptive statistics about the study areas are provided in Table A2.

Table A2: Descriptive statistics of study-areas

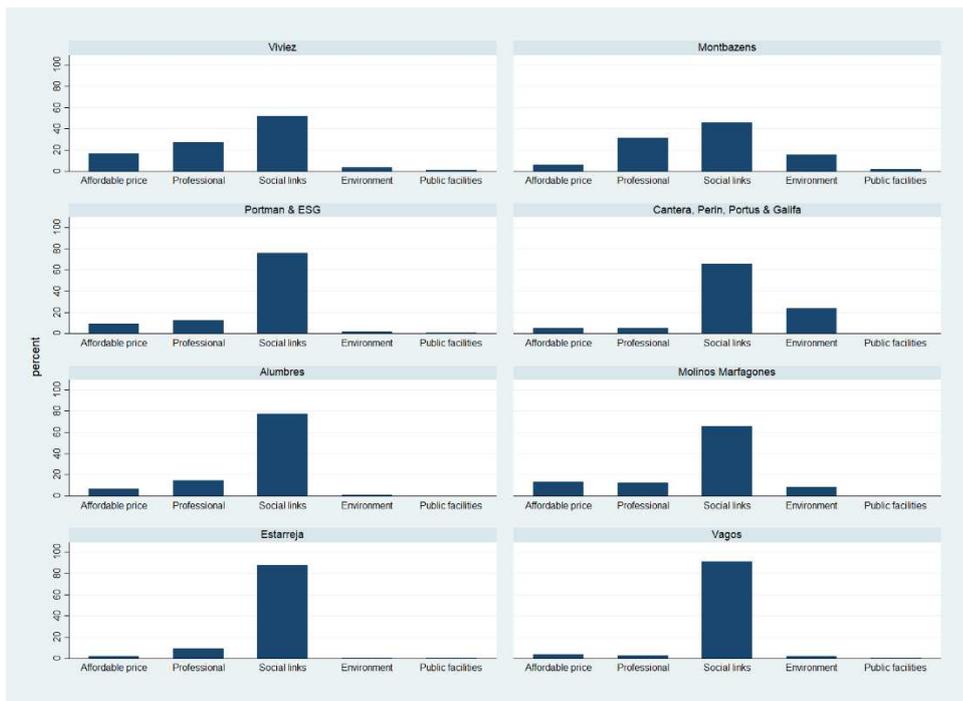
	Portman/ESG vs. control	Viviez vs. control	Alumbres vs. control	Estarreja vs. control
<i>Attachment to the area</i>				
Ratio of area lifetime/age (%)	0.11**	0.08*	0.30***	0,08***
Intention to move out in the next 5 years (proportion)	0,01	0.13***	0.10*	0,01
Living here for economic reasons (proportion)	0.19***	0.12**	-0.10*	-0,02
Living here for professional reasons (proportion)	0.31***	-0.10*	0.12**	0,20***
Living here for social/family reasons (proportion)	0.10*	-0,01	0.15**	0,01
Living here for environmental reasons (proportion)	-0.24***	-0.28***	-0.08**	-0,09***
Living here for public facilities (proportion)	0.04*	-0.07**	0,02	-0,02

Source: Authors' calculation from the CSPE database.

Figure A1 shows that “social ties” play an important role on residential motivations for all study-areas; as they are more often mentioned than “professional reasons” and “economic reasons”. Nonetheless, referring to Table A1, “professional reasons” are particularly mentioned as a motivation for living in Alumbres and Estarreja. Table A1 also shows that residents more often mention to have chosen to live in Viviez and Portman/ESG for “interesting economic reasons” (e.g. affordable housing price) compared to residents from non-polluted areas, who more often mention “environmental amenities” as residential motivation. This result is not surprising given the presence of an active industry supplying various jobs in both areas. Furthermore, strong community attachment and satisfaction are characteristics that are highlighted in several active industrial sites in Europe. However, perceived area attractiveness is significantly lower in these active sites (Viviez and Alumbres), compared to their respective control groups. This might be due to the presence of smokes, smells and noises.

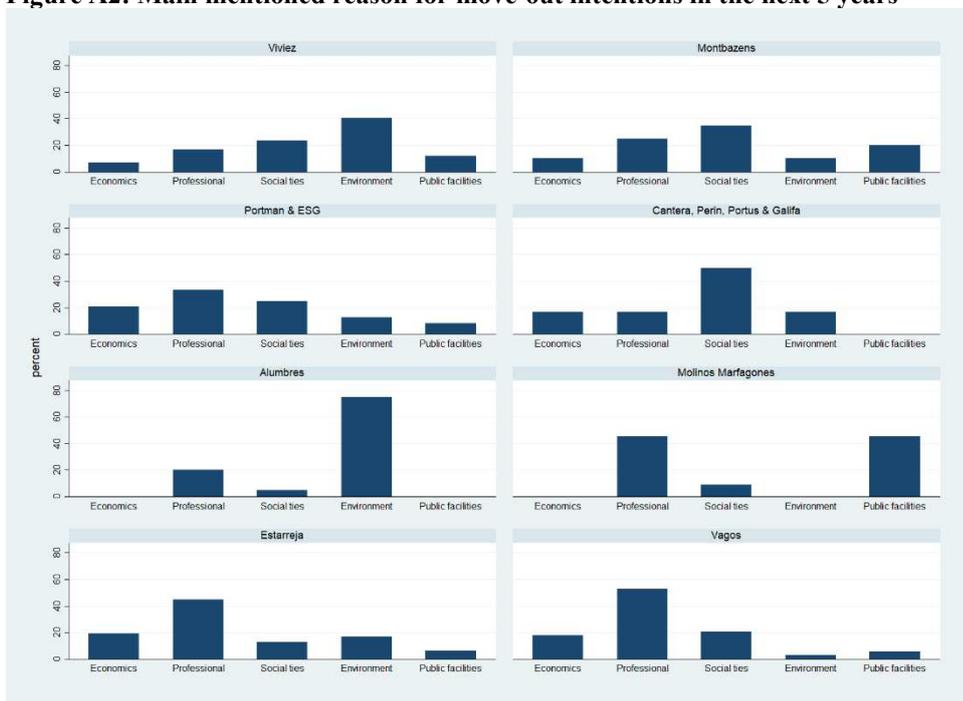
Table A1 also shows stronger intentions to move out for residents from Viviez and Alumbres compared to their respective control groups. In these two municipalities, “environmental issues” appear as the most mentioned motivation for moving out (Figure A2). In contrast, “environmental issues” does not appear as a specific move-out issue in Portman/ESG and Estarreja, compared to their respective counterparts. On the other hand, area attractiveness perceptions in Portman/ESG and Estarreja do not differ from their respective control areas. This absence of significant gap is not so surprising since, in both areas, there are close environmental amenities that residents can enjoy despite the presence of pollution (e.g. lagoon, sea, or ocean). In addition, pollution is not directly visible in Portman/ESG (i.e. mines closed since 1990) and can be relatively far in Estarreja for people living in other *freguesias* than Beduido & Veiros (where the chemical complex is located). In the same way, Table 1 does not show significant differences regarding move-out intentions in Portman/ESG and Estarreja, compared to their respective control group.

Figure A1: Main mentioned reason for living in the area



Source: CSPE database, authors' computation.

Figure A2: Main mentioned reason for move-out intentions in the next 5 years



Source: CSPE database, authors' computation.

Figure A3: Pollution perception in polluted and control areas as a function of household income

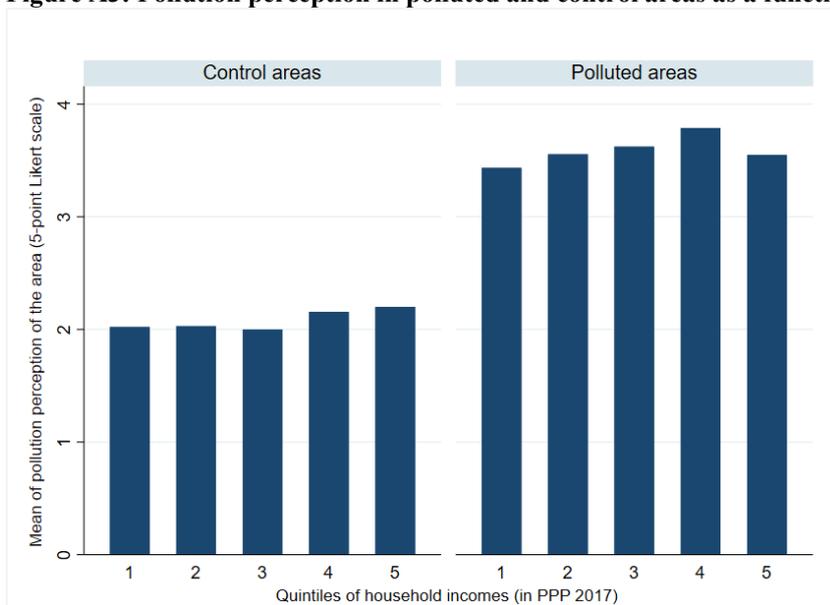
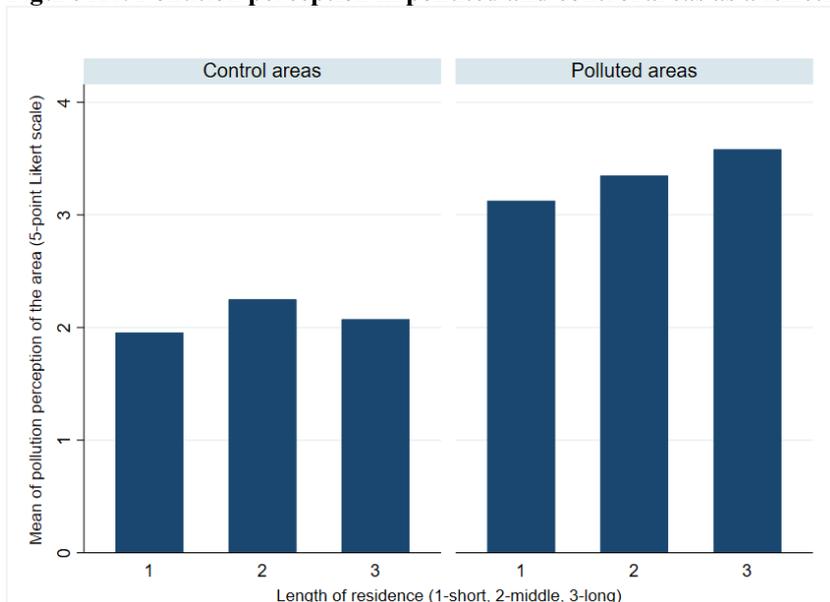


Figure A4: Pollution perception in polluted and control areas as a function of length of residence in the area



Figures A3 and A4 show that pollution perception is always higher in polluted areas than in control areas. Figure A3 underlines that pollution perception is increasing with income except for the highest income quantile in polluted areas. Figure A4 underlines that pollution perception in polluted areas is growing with the length of residence.

APPENDIX B: Additional materials

Table B1: Description of variables

Variable name	Description
Living in polluted areas (dummy)	Takes the value 1 if the household lives in a polluted area (Viviez, Sierra Minera, Estarreja) and 0 if it lives in a control area (Montbazens, Catargena West, Vagos).
Intention to move out in the next five years	Takes the value 1 if the household plans to move out of the community in the next five years, 0 otherwise.
Number of children (lower than 17yo)	Quantity of children in the household.
Number of young adults (18-29yo)	Quantity of young adults in the household.
Number of lower-middle age adults (30-44yo)	Quantity of lower-middle age adults in the household.
Number of higher middle age adults (45-64yo)	Quantity of higher-middle age adults in the household.
Number of old adults (higher than 65yo)	Quantity of old adults in the household.
Number of male members	Quantity of male members in the household.
Respondent is in a couple (dummy)	Takes the value 1 if the respondent is in a couple (married or free union), 0 otherwise.
Respondent is single (dummy)	Takes the value 1 if the respondent is single, 0 otherwise.
Number of members with foreign parents	Quantity of adult members with foreign parents in the household.
Number of rooms	Number of rooms in the housing.
Having a garden (dummy)	Takes the value 1 if the housing has a private garden, 0 otherwise.
Presence of family members around (dummy)	Takes the value 1 if there are family members living in the same community, 0 otherwise.
Lifetime in the area (in year)	Duration of residence in the community in years.
Willingness to live in a polluted land that decreases life expectancy by 5yo (1-to-5 scale)	The scale varies from 1 "a very low willingness to live in a place that may reduce life expectancy by 5yo" to 5 "a very high willingness to take this risk".
Number of members who obtained a high-school grade	Number of household members who obtained a high-school diploma
Wealth index (0-to-7 score)	Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0.
ln(monthly total household income)	Before being log-transformed, monthly total household incomes are corrected using purchasing power parities (PPP) based on 2017 US dollars: 0.581 for Portugal, 0.641 for Spain and 0.776 for France.
Respondent's height	Height of the respondent in meters
Respondent's parental education	Takes the value 1 if at least one parent of the respondent obtained a high-school diploma, 0 otherwise.

Source: Authors' computation from the CSPE database.

Table B2: Factors correlated to the probability of living in polluted areas (OLS estimates)

	(1)	(2)
Number of children (lower than 17yo)	-0.002 (-0.06)	-0.013 (-0.48)
Number of young adults (18-29yo)	0.021 (0.64)	0.026 (0.76)
Number of lower-middle adults (30-44yo)	-0.003 (-0.07)	0.011 (0.24)
Number of higher middle adults (45-64yo)	-0.039 (-1.49)	-0.043 (-1.06)
Number of old adults (higher than 65yo)	-0.061* (-2.06)	-0.079* (-1.94)
Number of male members	-0.024 (-1.03)	-0.007 (-0.22)
Respondent is in a couple (dummy)	0.036 (1.34)	0.073** (2.36)
Respondent is single (dummy)	0.074 (1.53)	0.119** (2.36)
Number of members with foreign parents	0.086*** (3.18)	0.080*** (3.00)
Number of rooms	0.044*** (3.89)	0.046*** (3.37)
Having a garden (dummy)	-0.161** (-2.69)	-0.146** (-2.33)
Presence of family members around (dummy)	0.156** (2.25)	0.174** (2.73)
Lifetime in the area (in year)	0.004*** (5.03)	0.004*** (5.37)
Willingness to live in a polluted land that decreases life expectancy by 5yo (1-to-5 scale)	0.084** (2.52)	0.080** (2.42)
Wealth index (0-to-7 score)	0.112*** (3.23)	
Square of wealth index	-0.017** (-2.73)	
ln(incomePPP)		1.083** (2.37)
Square of ln(incomePPP)		-0.073** (-2.44)
Constant	-0.119 (-0.47)	-4.052** (-2.37)
Country fixed effects	Yes	Yes
Observations	1,144	939
R-square	0.139	0.153

Notes: (1) Standard errors are robust to intra-group correlation. Significance levels are: ***1%, **5%, *10%,
(2) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0.
(3) Before being log-transformed, household incomes are corrected using purchasing power parities (PPP) based on 2017 US dollars: 0.581 for Portugal, 0.641 for Spain and 0.776 for France.

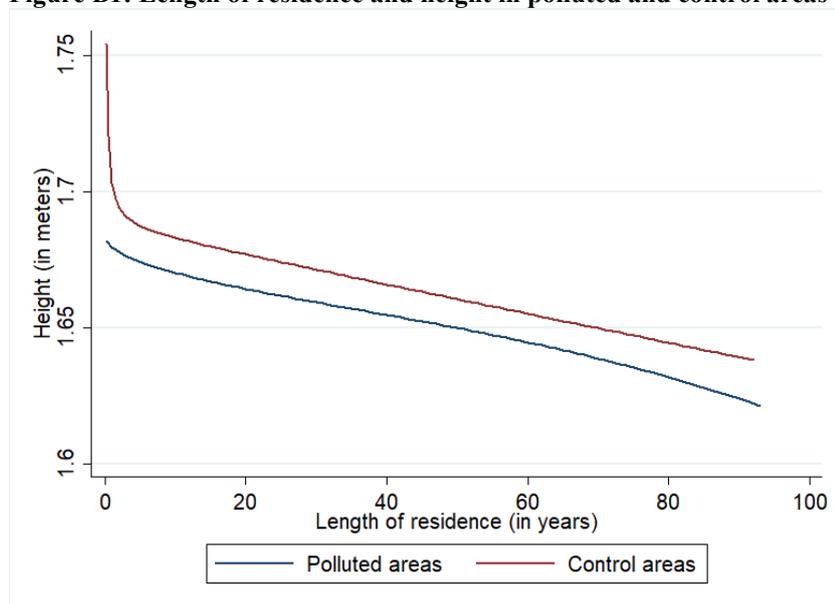
Source: Authors' calculation from the CSPE database.

Table B3: First-step estimates in IV model, linear regression of household socioeconomic status on instruments and control variables

	Number of educated members	Number of educated members	Wealth index	Wealth index	ln(income)	ln(income)
Height of the respondent (in meters)	1.252*** (4.09)	1.157*** (3.64)	1.611*** (3.57)	1.490*** (3.30)	0.943*** (4.54)	0.866*** (4.15)
Parental education of the respondent (at least a high-school diploma)		0.433*** (5.13)		0.333*** (3.11)		0.266*** (5.21)
Number of children (lower than 17yo)	0.006 (0.07)	-0.011 (-0.15)	0.055 (1.10)	0.029 (0.57)	0.036 (0.90)	0.027 (0.61)
Number of young adults (18-29yo)	0.607*** (9.86)	0.538*** (7.58)	0.248*** (3.74)	0.193*** (3.09)	0.267*** (5.50)	0.228*** (4.58)
Number of lower-middle age adults (30-44yo)	0.498*** (7.14)	0.440*** (7.88)	0.349*** (4.33)	0.310*** (3.72)	0.367*** (10.29)	0.350*** (9.59)
Number of higher middle age adults (45-64yo)	0.351*** (7.50)	0.343*** (9.30)	0.373*** (4.94)	0.365*** (4.47)	0.268*** (7.97)	0.275*** (8.33)
Number of old adults (higher than 65yo)	0.131** (2.13)	0.126** (2.36)	-0.044 (-0.42)	-0.041 (-0.41)	0.149*** (4.83)	0.170*** (5.21)
Number of male members	-0.082 (-1.50)	-0.061 (-0.99)	-0.028 (-0.47)	-0.017 (-0.24)	-0.040 (-1.49)	-0.031 (-1.04)
Respondent is in a couple (dummy)	0.070 (0.86)	0.042 (0.53)	0.696*** (7.67)	0.717*** (7.89)	0.252*** (6.79)	0.237*** (7.78)
Respondent is single (dummy)	-0.076 (-0.75)	-0.163 (-1.50)	0.324*** (3.05)	0.346*** (3.19)	0.089 (1.23)	0.059 (1.00)
Number of rooms	0.052*** (3.81)	0.047*** (3.18)	0.220*** (7.83)	0.235*** (9.60)	0.076*** (5.64)	0.071*** (5.19)
Constant	-2.536*** (-5.19)	-2.347*** (-4.45)	-0.787 (-1.03)	-0.703 (-0.91)	4.961*** (13.73)	5.085*** (13.71)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,147	1,098	1,147	1,098	943	911
R-square	0.302	0.329	0.340	0.353	0.436	0.464

Notes: (1) Standard errors are robust to intra-group correlation. Significance levels are: ***1%, **5%, *10%,
(2) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0. Incomes are in US\$ PPA.
Source: Authors' calculation from the CSPE database.

Figure B1: Length of residence and height in polluted and control areas



Note: Fractional polynomial estimates are used.
Source: Authors' calculation from the CSPE database.

Table B4: Reduced-form regression of the dependent variable on the instruments

	Prob(LivingPollutedAreas)
Height of the respondent (in meters)	-0.548*** (-3.02)
Parental education of the respondent (at least a high-school level)	-0.075* (-1.80)
Number of children (lower than 17yo)	-0.003 (-0.10)
Number of young adults (18-29yo)	0.019 (0.56)
Number of lower-middle age adults (30-44yo)	-0.003 (-0.07)
Number of higher middle age adults (45-64yo)	-0.061* (-1.84)
Number of old adults (higher than 65yo)	-0.037 (-1.04)
Number of male members	-0.018 (-0.59)
Respondent is in a couple (dummy)	0.049 (1.08)
Respondent is single (dummy)	0.085 (1.64)
Number of rooms	0.045*** (4.51)
Country fixed effects	Yes
Observations	1,098
R-square	0.0295

Source: Authors' calculation from the CSPE database.

Table B5: Factors correlated to the intention to move out in the next 5 years, OLS estimates (full table of Eq.3)

	(1)	(2)	(3)	(4)	(5)
Average age of adults (in years)	-0.002** (-2.47)	-0.003** (-2.55)	-0.002** (-2.56)	-0.000 (-0.35)	-0.000 (-0.43)
Number of male members	-0.010 (-0.78)	-0.011 (-0.90)	-0.001 (-0.06)	-0.001 (-0.06)	0.001 (0.07)
Respondent is in a couple (dummy)	-0.004 (-0.30)	-0.003 (-0.18)	0.002 (0.09)	0.002 (0.10)	0.001 (0.04)
Respondent is single (dummy)	0.056** (2.25)	0.056** (2.34)	0.067** (2.48)	0.067** (2.45)	0.068** (2.57)
Number of members with foreign parents	0.027 (0.99)	0.026 (0.96)	0.036 (0.84)	0.039 (0.93)	0.036 (0.88)
Number of rooms	-0.004 (-0.59)	-0.003 (-0.61)	-0.006 (-0.83)	-0.006 (-0.79)	-0.006 (-0.76)
Having a garden (dummy)	-0.040 (-1.12)	-0.037 (-1.03)	-0.023 (-0.76)	-0.024 (-0.81)	-0.022 (-0.81)
Presence of family members around	-0.043* (-2.06)	-0.045** (-2.14)	-0.037 (-1.32)	-0.037 (-1.34)	-0.038 (-1.44)
Lifetime in the area (in year)	-0.001** (-2.67)	-0.001*** (-2.84)	-0.001*** (-3.44)	-0.001*** (-3.27)	-0.001*** (-3.10)
Willingness to live in a polluted land that decreases life expectancy by 5yo (1-to-5 scale)	0.007 (0.62)	0.007 (0.55)	0.001 (0.15)	0.001 (0.12)	-0.002 (-0.17)
Area attractiveness index (3-to-15 scale)	-0.010*** (-3.68)	-0.009*** (-3.45)	-0.010*** (-3.55)	-0.010*** (-3.39)	-0.009*** (-3.07)
Perceived knowledge about anti-pollution protection methods (1-to-5 scale)	-0.012* (-1.90)	-0.011 (-1.63)	-0.011 (-1.60)	-0.011 (-1.58)	-0.011 (-1.64)
Associative involvement (dummy)	0.031 (1.04)	0.033 (1.16)	0.028 (0.97)	0.025 (0.85)	-0.036 (-0.74)
Social cohesion perception (1-to-5 scale)	-0.022** (-2.34)	-0.024** (-2.51)	-0.022* (-1.79)	-0.024* (-1.92)	-0.026** (-2.18)
Regular participation in community parties (dummy)	-0.022 (-0.99)	-0.022 (-0.97)	-0.013 (-0.54)	-0.016 (-0.64)	-0.057** (-2.59)
Number of members who obtained a high school grade	-0.006 (-0.75)	-0.002 (-0.26)	-0.016 (-1.64)	-0.014 (-1.43)	-0.015 (-1.70)
Wealth index (0-to-7 score)		-0.019** (-2.35)			
Income Quintile 1			0.025 (0.47)	-0.002 (-0.05)	-0.022 (-0.44)
Income Quintile 2			0.007 (0.13)	-0.005 (-0.11)	-0.016 (-0.30)
Income Quintile 3			0.020 (0.28)	0.022 (0.33)	0.017 (0.25)
Income Quintile 4			-0.004 (-0.11)	-0.009 (-0.26)	-0.019 (-0.55)
Polluted area (dummy)	0.034* (1.90)	-0.045 (-1.09)	0.139** (2.13)	0.292*** (3.18)	0.229** (2.47)
Wealth*PollutedArea		0.022* (1.93)			
IncomeQ1*PollutedArea			-0.105 (-1.67)	-0.058 (-1.03)	-0.039 (-0.68)
IncomeQ2*PollutedArea			-0.078 (-1.18)	-0.053 (-0.85)	-0.044 (-0.70)
IncomeQ3*PollutedArea			-0.166* (-2.02)	-0.165** (-2.09)	-0.166** (-2.12)
IncomeQ4*PollutedArea			-0.136** (-2.40)	-0.130** (-2.44)	-0.124** (-2.43)
AverageAge*PollutedArea				-0.003*** (-3.55)	-0.003*** (-3.50)
Association*PollutedArea					0.073* (1.65)
CommunityParties*PollutedArea					0.113** (2.29)
Constant	0.514*** (5.37)	0.585*** (5.31)	0.499*** (5.11)	0.411*** (4.16)	0.449*** (4.70)
Country fixed-effect	Yes	Yes	Yes	Yes	Yes
Observations	1,099	1,099	900	900	900
R-square	0.110	0.114	0.133	0.139	0.150

Notes: (1) Standard errors are robust to intra-group correlation. Significance levels are: ***1%, **5%, *10%,
(2) Wealth index is the sum of the following owned (or not) assets: former house, second house, car, air conditioner, computer, cellphone and financial assets. Thus, the wealthiest households have a score of 7 while the most deprived household a score of 0.
(3) Area attractiveness index is the sum of the following 1-to-5 perception scores: global attractiveness, availability of public services and availability of shops and retails. Thus, areas perceived as the most attractive have a score equal to 15, while areas perceived as the most deprived have a score of 3.

Source: Authors' calculation from the CSPE database.

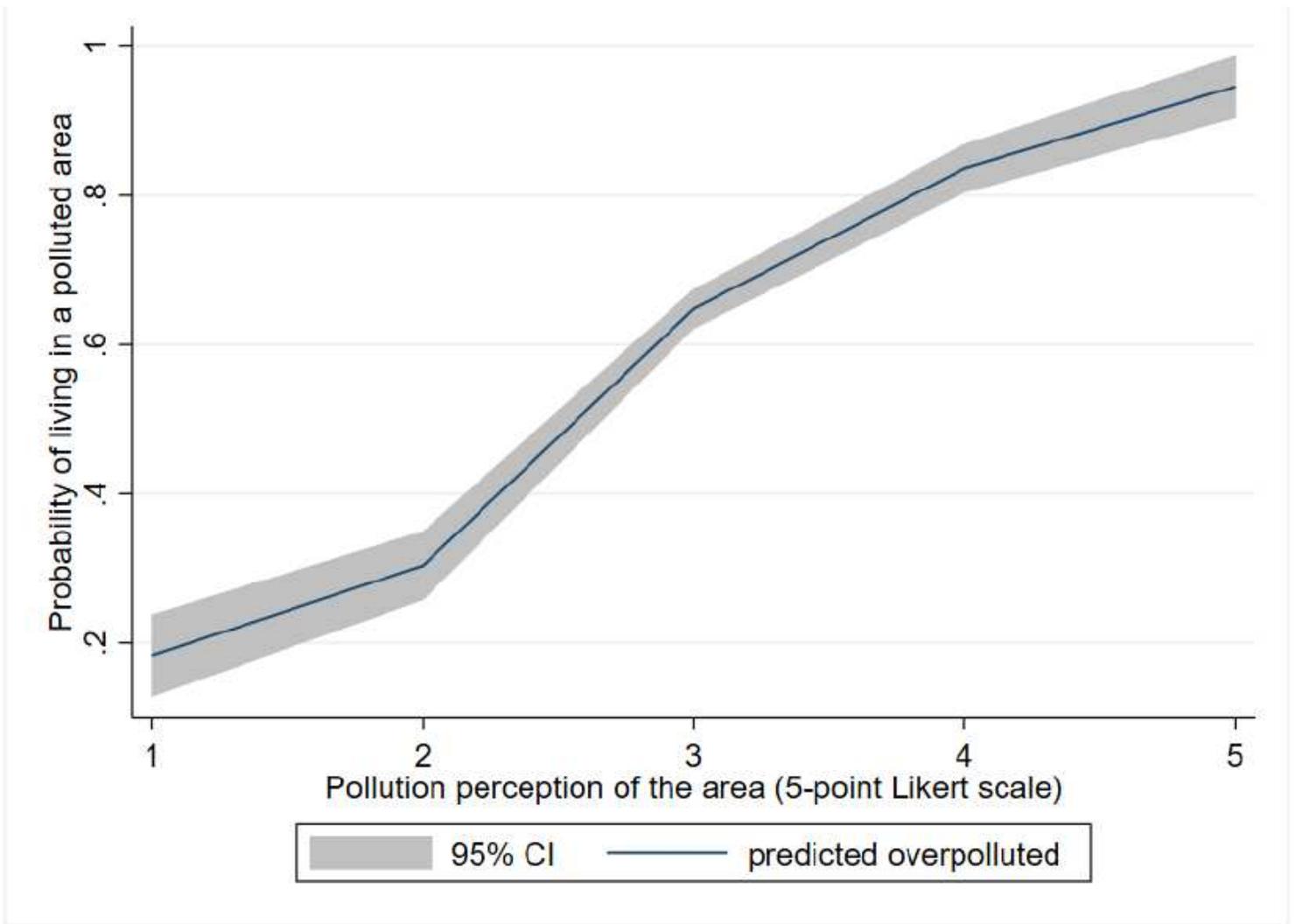
Figures



Source: OSM, authors' computation.

Figure 1

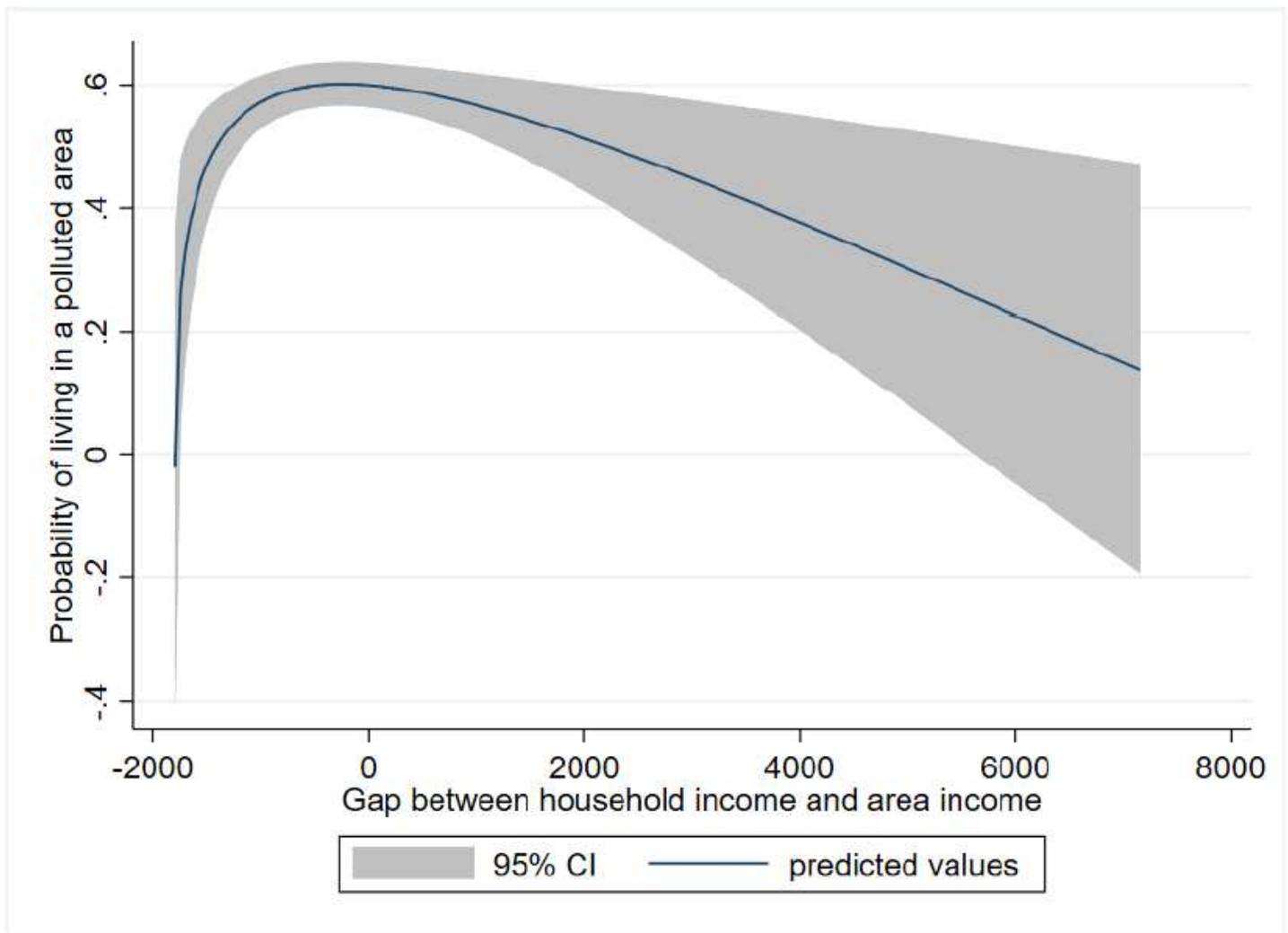
Mapping of polluted and control areas. Source: OSM, authors' computation.



Note: Fractional polynomial estimates are used.
Source: Authors' calculation from the CSPE database.

Figure 2

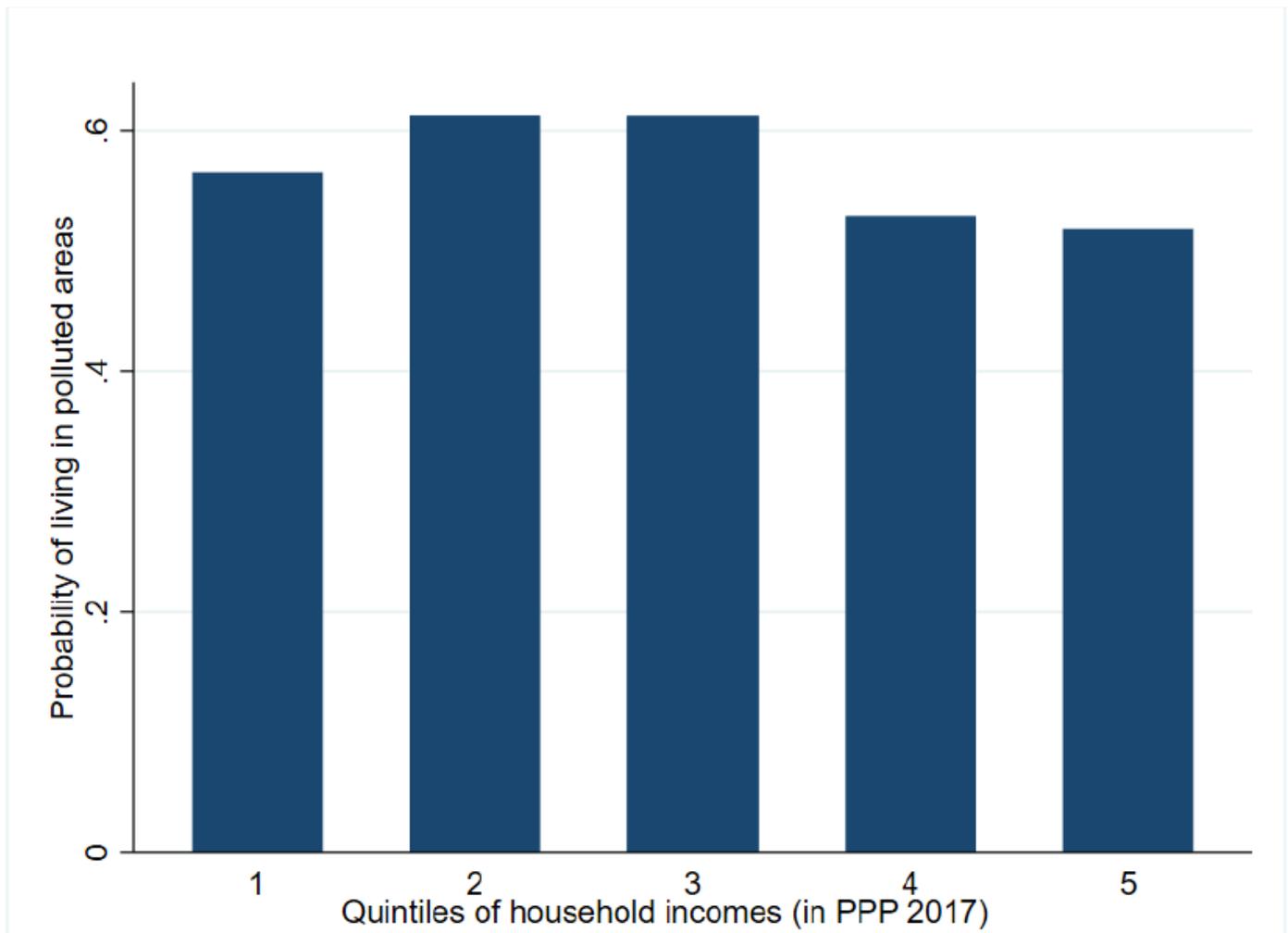
Respondent's pollution perception and probability of living in polluted areas. Note: Fractional polynomial estimates are used. Source: Authors' calculation from the CSPE database.



Note: Fractional polynomial estimates are used.
 Source: Authors' calculation from the CSPE database.

Figure 3

Income gaps in polluted areas. Note: Fractional polynomial estimates are used. Source: Authors' calculation from the CSPE database.



Source: Authors' calculation from the CSPE database.

Figure 4

Household incomes (in \$PPP) and probability of living in polluted areas. Source: Authors' calculation from the CSPE database.