

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeeem

The effect of air pollution on fertility in 657 European regions

 Árpád Stump^a, Ágnes Szabó-Morvai^{b,c} ^{*,*}
^a Ministry of Finance, 1051 Budapest, József nádor tér 2-4, Hungary^b HUN-REN KRTK KTI, 1097 Budapest, Tóth Kálmán u. 4, Hungary^c University of Debrecen, 4032 Debrecen, Egyetem tér 1, Hungary

ARTICLE INFO

JEL classification:

Q53

J13

I14

Keywords:

Ambient air pollution

Fertility

Instrumental variables

Particulate matter

ABSTRACT

This study examines the impact of ambient air pollution on birth rates in Europe. We estimate the causal effect of air pollution on fertility by utilizing variations in wind speed and the number of heating days as instrumental variables for air quality. Our analysis encompasses 657 regions of NUTS 3 level, with each region having 2 to 6 years of observations between 2013 and 2020. Thus, our study is the first to extend this analysis to multiple countries, pollutants, and years. Our findings indicate that a one standard deviation increase in particulate matter concentration levels leads to a 14.1% decrease in birth rates the following year and an additional 17.2% decrease two years later. Moreover, a similar increase in air pollution has a more pronounced adverse effect on fertility in countries with lower GDP. Other pollutants play little role in shaping fertility outcomes. This result is important for environmental policies with limited resources.

1. Introduction

In the past 70 years, fertility rates have been falling in most developing countries, raising concerns about the sustainability of pension and healthcare systems. At the same time, air pollution has become an important environmental and health problem throughout the world (Fowler et al., 2020). In this research, we examine whether and how much air pollution affects fertility rates. The results of the analysis have direct policy relevance for developed countries aiming to combat low fertility.

Air pollution is a significant health concern in modern societies. The World Health Organization (WHO) ranked ambient air pollution among the ten most important threats to global health in 2019 (WHO, 2019). According to the 2018 Special Report of the European Court of Auditors (ECA, 2018), the number of years of healthy life lost due to ambient air pollution is on average 0.75 per 100 inhabitants of Europe. In 2021, 97% of the urban population in Europe was exposed to concentrations of particulate matter that exceeded WHO guidelines (European Council, 2024). More specifically for this study, air pollution is significantly correlated with fertility, as shown by several studies summarized by the meta-studies of Frutos et al. (2015), Conforti et al. (2018), Jahnke et al. (2022), and Siegel et al. (2023). These meta-analyses present several studies that examine the associations of air pollution concentrations with fertility, fecundability, sperm count, miscarriages, and stillbirths. Many studies examine only one pollutant at a time, which likely biases estimates, as the excluded pollutants may be correlated with both the included pollutant and fertility.

Moreover, some studies include multiple pollutants (e.g., Slama et al., 2013; Nieuwenhuijsen et al., 2014; Nobles et al., 2018; Legro et al., 2010; Green et al., 2009) to avoid biased estimates due to omitted pollutants. Still, these studies acknowledge that they measure associations between air pollution and fertility-related outcomes, as confounders such as industrial activities and traffic may still be omitted from these regressions. Recent research cautions against drawing policy conclusions from associational

* Corresponding author at: HUN-REN KRTK KTI, 1097 Budapest, Tóth Kálmán u. 4, Hungary.

E-mail address: szabomorvai.agnes@krtk.hun-ren.hu (Á. Szabó-Morvai).

<https://doi.org/10.1016/j.jeeem.2024.103111>

Received 27 April 2024

Available online 20 January 2025

0095-0696/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

estimates (Currie and Walker, 2011; Godzinski and Suarez Castillo, 2021); however, conducting causal estimations remains challenging.

In contrast, our quasi-experimental study measures a causal relationship, similar to Godzinski and Suarez Castillo (2021), which examines the effect of air pollution on morbidity and mortality. We employ a multiple pollutant method, including all pollutants measured by European countries in our analysis. This approach is combined with an instrumental variable strategy, using wind speed and the number of heating days as instruments for pollution concentrations. We analyze yearly birth rates and air pollution data from 36 countries in Europe and neighboring regions at the NUTS 3 region level,¹ incorporating concentration levels of the ten most significant pollutants. To our knowledge, this is the first quasi-experimental study to examine the effects of air pollution on fertility across a large number of countries and years while including multiple pollutants. Consequently, our findings offer strong internal and external validity, representing a significant advancement in this area of research.

Moreover, a substantial body of literature thoroughly establishes the mechanisms by which various pollutants can affect fertility rates. However, these studies do not assess the overall magnitude of the problem or identify which pollutants have the most significant effects on fertility. For instance, Levine et al. (2017) documents a dramatic worldwide decrease in sperm count—over 50% between 1973 and 2011—and there is direct evidence of a causal relationship between air pollution and declining semen quality, including concentration, count, and motility (Qian et al., 2022). Additionally, particles can induce inflammatory processes and hormonal disruptions that negatively impact fecundability (Siegel et al., 2023). Recent evidence also suggests that particulate matter may reach the placenta, increasing risks to the fetus (Bové et al., 2019). Pollutants have also been shown to negatively affect ovarian reserve and increase the risk of conditions such as endometriosis and polycystic ovary syndrome (PCOS), which reduce fecundability (Siegel et al., 2023).

Our study contributes to a broader body of quasi-experimental research on the various health effects of air pollution. Numerous studies demonstrate that higher levels of pollutants such as PM_{2.5}, PM₁₀, and ozone negatively impact child health outcomes, increasing the likelihood of low birth weight, infant mortality, and asthma incidence (Alexander and Schwandt, 2022; Lleras-Muney, 2010; Gutierrez, 2015; Chay and Greenstone, 2003). For example, Currie and Schwandt (2016) provide evidence that exposure to the toxic dust from the 9/11 dust cloud significantly increased the likelihood of preterm birth and low birth weight.

Additionally, this study contributes to the broader literature on fertility. Previous research has identified various factors that influence fertility rates, both through the demand for children and through maternal and paternal health. Factors affecting the demand for children include labor market shocks, availability of child subsidies, and childcare coverage, among others (see Doepke et al., 2023; Alderotti et al., 2021, for a review). Factors influencing parental health include the affordability of the healthcare system, environmental factors, and lifestyle choices, among many others (see systematic reviews such as Campbell et al., 2015; Oldereid et al., 2018).

In the next section, we discuss our data collection. Section 3 describes our empirical method and provides details of the robustness checks. We present our main results and heterogeneous effects, and the related robustness checks in Section 4. In Section 5, we present the results of linear simulation results using our main model. Finally, Section 6 concludes the article.

2. Data

First, we collect air quality data from the European Environment Agency (EEA) using a web scraping technique in order to gather the air quality data collected using a representative sample of measuring stations that member states upload to the Internet. Note that the most frequently studied pollutants in the literature are particulate matter, carbon-monoxide, sulfur-dioxide, and nitrogen-dioxide. We extend the list of pollutants and collect information about nitrogen dioxide (NO₂), nitrogen monoxide (NO), different nitrogen oxides (NO_x), ozone (O₃), sulfur dioxide (SO₂), different sizes of particulate matter (PM_{2.5} and PM₁₀), benzene (C₆H₆), lead (Pb), and carbon monoxide (CO). We downloaded more than 1.1 billion data points (see Table A.5 in the Appendix). CO pollution is measured in mg/m³, while all the other pollutants are measured in µg/m³.

We first clean the database and delete observations that are not on hourly or daily frequency, observations with negative concentration values, and all non-validated observations (mostly missing values). Second, we calculate daily averages from the hourly data. We connect the stations to NUTS 3 regions using the GPS coordinates of the measurement stations. Since countries only report a representative selection of their air quality data to the EEA, we do not have stations in each NUTS 3 region. When there are more stations in a region, we calculate the average daily concentration across stations for each NUTS 3 region.

We aggregate the daily NUTS 3 average concentration levels to the yearly frequency using three methods. First, we calculate the yearly mean for each pollutant and for each NUTS 3 region. Second, we calculate deciles of the daily pollution levels across every NUTS 3 region for the whole observation period. Then for each year and NUTS 3 region, we count the number of days when the pollution concentration was in or above the given decile. For instance, D₉_{rt_p} shows the number of days when the concentration level of pollutant *p* was in the 9th or the 10th pollution decile in region *r* and year *t*. Third, we also examine the concentrations relative to the European Air Quality Standards as of 2023 (AQS, see Table A.6 and Figures A.7 and A.8 in the Appendix).² For example, AQS 125%_{rt_p} is the number of days when the concentration of pollutant *p* exceeded 125% of the relevant concentration limit in year *t* and region *r*.

¹ The European Union has developed a region classification system known as the Nomenclature of Territorial Units for Statistics (NUTS). NUTS 3 refers to small regions, such as districts (Kreis) in Germany or prefectures in Greece; see NUTS (2024, for details).

² For some pollutants, such as NO and NO_x, no daily mean pollution threshold values are set by the EU. For these pollutants, we use the annual target value or the maximum daily 8-hour mean value.

Birth rates are based on Eurostat data (EUROSTAT, 2022), and are calculated as the ratio between the number of live births and the number of women of reproductive age (15–44) on 1st January. The total female population in the NUTS 3 regions is used as weights in the regressions. We also use EUROSTAT data to include NUTS 3-level GDP per person.

The NUTS 3 level yearly heating degree days (HDD) data are provided by the Joint Research Centre’s AGRI4CAST Resources Portal (EUROSTAT, 2021). HDD is a weather-based technical index which is higher if there is more need for heating, taking into account the outdoor temperature, the usual indoor temperature, and technical details of the buildings.

We also use NUTS-2 level daily wind speed data (measured in km/h) from the Copernicus Climate Change Service (European Commission, 2020). From the daily observations, we calculate the yearly mean wind speed.

Tables A.7 and A.8 in the Appendix show the overall coverage of the variables. Note that we only include the NUTS 3 regions that have at least one pollution data observation. HDD has good coverage, as there is data available for every NUTS 3 region in the EU, but, in general, there are no observations for countries outside of the EU (e.g., the UK). The birth rate is available for most of the regions, however, demography structure indicators (e.g., the female population aged 15–44) were only available from 2014 and on.

As Tables A.7 in the Appendix indicates, there is a trade-off between using many pollutants and many regions in our regressions. As a solution, we include pollutants with a low number of observations (NO, C₆H₆, and Pb) in the LASSO regressions, but we omit them from the main regression analyses. Still, there is a strong correlation between the yearly levels of the remaining pollutants (Table A.9 in the Appendix). For example, the correlation coefficient between PM_{2.5} and PM₁₀ pollution levels is 0.81. NO₂ has a very strong correlation with NO_x, the correlation coefficient is 0.74.

To circumvent this issue, we use principal factor analysis to combine the highly correlated pollutants. As a result, we are left with three pollutant variables in the main regressions: *PM Factor*, *NO Factor*, and *SO₂*. *PM Factor* includes PM₁₀, PM_{2.5}, and CO. The primary sources of particulate matter pollution are local combustion (e.g., traffic, metal industry plants), residential heating with solid fuels in cold seasons and biological material (e.g., vegetative debris, spores and pollen) in warm and dry seasons (Sillanpää et al., 2006). The largest part of CO emissions come from the incomplete combustion of vehicle fuels. *NO Factor* includes NO₂, NO_x stemming from the combustion of fossil fuels, like car emissions and O₃ which is produced as a result of a chemical reaction of NO₂ and NO_x with oxygen in the presence of heat and sunlight. *SO₂* is included in the regressions by itself. This pollutant is mainly created when electric utilities and power plants burn coal and oil. For the details of the factorization, see Table A.10 in the Appendix.

3. Empirical method

3.1. Naive OLS

Our goal is to estimate the effect of air pollution on fertility. First, we estimate naive regressions of pollution indicators in the previous year on the natural logarithm of the birth rate. In our main specification, the mean concentrations of ambient air pollution is included. It is important that all pollutants are included in the regression at the same time because they are correlated and many of them may affect fertility. Examining only one pollutant at a time would likely cause the estimates to suffer from omitted variables bias.

The observations are aggregated to the year (t) and region (r) level. We include year fixed effects (η_t) to control for any general shock that affected the regions at the same time, such as Europe-wide economic cycles. We also include region fixed effects (λ_r) to control for unobserved differences between regions that remain unchanged over a few years, such as social norms that influence environmental awareness and fertility preferences. Finally, we allow for region-specific linear time trends ($\lambda_r \times t$) of fertility in the model. Throughout the analysis, we use robust standard errors, clustered at the regional level. We estimate the following model:

$$\ln(Y_{rt}) = \sum_{i=1}^5 \beta_i P_{rt-1}^i + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{1}$$

where Y_{rt} is the birth rate, the number of births per 1000 women of age 15 to 44 in region r and year t, P_{rt}^i is the concentration level of pollutant i in region r and year t, and ε_{rt} is the error term. We calculate robust standard errors clustered at the regional level.

Air pollution can possibly also affect fertility in the long run. To test this, we include 2-year lags of the pollutants in our second specification.

$$\ln(Y_{rt}) = \sum_{i=1}^5 \beta_i P_{rt-1}^i + \sum_{i=1}^5 \gamma_i P_{rt-2}^i + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{2}$$

Next, we include regional-level GDP as an additional control variable.

$$\ln(Y_{rt}) = \sum_{i=1}^5 \beta_i P_{rt-1}^i + \sum_{i=1}^5 \gamma_i P_{rt-2}^i + \tau GDP_{rt-1} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{3}$$

3.2. Instrumental variables approach

There can be other region-specific time-variant variables that we cannot observe, such as future expectations or regional variations in spending on public services (health services and public transport). Not controlling for them in the analysis may lead to a bias of unknown direction and size in our point estimates.

To avoid this source of bias, we follow an instrumental variables design. Our instruments are wind speed and the number of heating days. These variables have previously been used as instruments for pollution in the literature. Knittel et al. (2016) use local weather conditions, Schwartz et al. (2015), Schwartz et al. (2017), Zabrocki et al. (2022), Godzinski and Suarez Castillo (2021), and Deryugina et al. (2019) use wind direction and speed, and Arceo et al. (2016) use temperature (thermal inversions) to instrument endogenous ambient air pollution concentrations.

In the bulk of the previous literature, only one or just a few pollutants have been included in the regressions. In this case, even using an instrumental variable design does not provide unbiased point estimates, because the exclusion restriction is likely not valid when the instrument affects the pollutants omitted from the regressions, which in turn may affect the dependent variable (Benmarhnia et al., n.d.).

We can use wind speed and the number of heating days and their nonlinear functions as instruments because they affect ambient air pollution concentration and composition. Higher wind speed helps to dissipate high concentrations of ambient air pollution. However, on cold winter days, the emissions increase as a result of the heating activity. The number of heating days captures this relation. Figure A.6 in the Appendix depicts the associations between the pollutants and the instruments.

The instrumental variables strategy provides unbiased estimates if the exclusion restriction holds. This ultimately consists of two parts. First, instrumental variables should be exogenous to fertility rates in the sense that these are not affected by any other factors that may correlate with fertility rates, such as economic cycles. In the case of wind and weather, it is safe to assume that the daily and short-term yearly deviations from the average are not affected by any of these factors.

Second, it is important that these weather conditions only affect fertility through air pollution and no other channels. In the previous literature, we know of no evidence that wind speed or the number of cold days would directly affect fertility rates (see e.g. Lam and Miron, 1996). The previous literature has well-documented the relationship between hot weather and fertility (Hajdu and Hajdu, 2021; Barreca et al., 2018), but this relationship does not appear to extend to cold weather.

A potential concern with using the number of heating days as an instrument is that it may influence fertility not through temperature but through the maternal characteristics prevailing in winter: previous literature has established that fertility and maternal characteristics vary by season. Buckles and Hungerman (2013) find that mothers of children born in winter are more likely to be teenagers and have a high school diploma. Moreover, they find that their results are driven by wanted pregnancies, which indicates that it is not the actual cold weather in winter (“conditions at conception have almost no explanatory power”), rather the “expected weather at birth are driving” the results. This points to the importance of preferences for certain calendar seasons by mothers of various background characteristics. We avoid this problem by using annual data, thereby eliminating all seasonal correlations from our variables, as the number of winter days remains constant across years and regions (we are not comparing winter to spring, but year to year).

Also, it might be a concern that a high number of cold days could affect agricultural activities and thereby fertility negatively. Nevertheless, the number of heating days occurs mostly in winter, so it has a moderate effect on agriculture. In addition, the share of agriculture in GDP is very low in the EU, 1.6% on average in 2021 (WorldBank, 2024). Thus, it is unlikely that the number of heating days could have a meaningful effect on fertility through employment or GDP. Likewise, as the food markets of the EU are integrated, it is not likely that cold days in one region could significantly affect food prices and through them fertility.

As mentioned before, we include 3 pollution variables (PM Factor, NO Factor, and SO₂) in the main regressions, thus we need at least 3 instruments. As Figure A.6 shows, there is a non-linear relationship between pollutants and instruments. Thus, we use non-linear combinations of two instruments, including squared and cubic values, interactions and indicator functions, altogether 22 instruments³: Mean wind speed (WS), WS^2 , WS^3 , Number of heating days (HDD), HDD^2 , HDD^3 , $HDD \times WS$, $HDD \times WS^2$, $HDD \times WS^3$, $HDD \times WS$, $HDD^2 \times WS$, $HDD^3 \times WS$, $Days(WS > 4 \text{ km/h})$, $Days(WS > 5 \text{ km/h})$, $Days(WS > 6 \text{ km/h})$, $Days(WS > 7 \text{ km/h})$, $Days(WS > 8 \text{ km/h})$, $[Days(WS > 4 \text{ km/h})]^2$, $[Days(WS > 5 \text{ km/h})]^2$, $[Days(WS > 6 \text{ km/h})]^2$, $[Days(WS > 7 \text{ km/h})]^2$, $[Days(WS > 8 \text{ km/h})]^2$. We include the same lagged values of the instruments as of the pollutants. In our main specification, one and two-year lags of the pollutants and the instruments are included. We run two-stage least squares (2SLS) regressions. The first-stage results show how strong and significant the relationship is between the instruments and pollution concentrations. The first stage for the pollution concentrations one year before birth is:

$$P_{r,t-1}^i = \sum_{j=1}^2 \sum_{k=1}^{22} (\pi_{k,t-j} Z_{k,t-j}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{4}$$

where subscript j denotes the number of lags, and k is the kth instrument from the list of instruments, r denotes region, and t stands for year. The first stage for the pollutants two years before birth is:

$$P_{r,t-2}^i = \sum_{j=1}^2 \sum_{k=1}^{22} (\pi_{k,t-j} Z_{k,t-j}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{5}$$

³ We choose a suitable subset from the following instrument list.

The reduced-form equations are the following:

$$\ln(Y_{rt}) = \sum_{j=1}^2 \sum_{k=1}^{22} (\gamma_{k,t-j} Z_{k,t-j}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{6}$$

3.3. Robustness checks

First, we include other measures of ambient air pollution concentrations. Besides the mean pollution concentration, we measure the pollution concentration with number of days when the concentrations exceeded 75% of the European air quality standard concentration limits and the number of days when the concentration levels reached the 10th decile of pollution, as described in Section 2.

Second, we use different factorization methods to generate PM Factor and NO Factor variables. In the baseline specification, we use the principal factor method, and as a robustness check, we use the principal-component factor method, iterated principal-factor method, and maximum-likelihood factor method.

Third, we check whether the results of the LASSO estimation vary with the specification. In the baseline result, we use the minimum of the CV function to select λ . In robustness checks, we use the “one-standard-error rule” (Hastie et al., 2015); the minimum value of the BIC function; the minimum of the BIC function where models are fit for all lambdas in the grid until the tolerance value is reached; adaptive ridge (adaptive lasso, using the ridge estimator to construct the initial weights in the first lasso); adaptive steps (adaptive lasso with 100 lassos); and adaptive power 1.5 (adaptive lasso, where weights are raised to the 1.5th power). In addition, we check alternative seeds.

Fourth, we plot mean pollution values and birth rates by year and run regressions without COVID-19 years.

Fifth, we check whether our results are robust to the choice of instruments.

3.4. Heterogeneity

In our dataset, we included many EU regions, and thus are able to present a heterogeneity analysis. We divide the sample by the average levels of PM concentrations throughout the observation period. The high pollution subsample includes regions with higher than median PM pollution levels, and the low pollution subsample includes those with lower than median levels. Next, we do the same with GDP and run the 2SLS regressions on these subsamples. These two dimensions are somewhat correlated ($\rho = -0.3$), as the wealthier regions are less polluted, in general. Still, about 30% of the regions are in the “high pollution - high GDP” or “low pollution - low GDP” categories.

3.5. LASSO estimations

In the OLS and 2SLS regressions, we simplified the estimations in two ways. First, we omitted NO, C₆H₆, and Pb due to the low number of observations. If these pollutants affect fertility, our regressions suffer from omitted variables bias. Second, we combined individual pollutants into factors, but we may want to know how much each of these affects fertility.

We use the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996) which allows us to include each pollutant that we observe in the data. Thus, we can get an idea about how important those factors are that we omitted from the main specifications, and whether we need to worry about them. Moreover, LASSO also permits us to separately evaluate the importance of each pollutant in the same regression, without combining them in factor variables.

LASSO is very similar to ordinary least squares (OLS) regression, except that the minimand function of the optimization does not only include the residual sum of squares (RSS), but also a penalty term (λ) that increases with larger absolute values of the regression coefficients (see Eq. (7)). In practice, this optimization method finds the curve that fits the data best, using as small a number of variables with $\beta \neq 0$ as possible.

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| = RSS + \lambda \sum_{j=1}^p |\beta_j| \tag{7}$$

In other words, the lasso technique uses shrinkage and thus offers a simple way to select a model with reasonably few variables, which performs the best out-of-sample prediction of the dependent variable (James et al., 2013). Recently, it has become an accepted method to use machine learning techniques to select control variables, see Angrist and Frandsen (2019), Böheim and Stölinger (2020), and Fluchtmann et al. (2020), for example. In the baseline LASSO specification, we use the cross-validation function to select λ and we use a linear LASSO model.

Table 1
Summary statistics - mean pollution concentrations.

	Mean	SD	Min	Max	N
Birth rate	53.51	10.98	21.67	167.41	7842
PM ₁₀	18.76	13.93	0.00	137.63	9392
PM _{2.5}	7.99	8.18	0.00	85.00	9392
CO	0.17	0.27	0.00	4.23	9392
SO ₂	2.70	5.14	0.00	157.44	9392
NO ₂	15.96	11.94	0.00	118.20	9392
NOx	19.80	26.04	0.00	214.39	9392
O ₃	38.78	24.54	0.00	111.56	9392
Wind speed	3.03	0.88	1.05	5.72	9200
Heating Degree Days	2620.17	867.19	266.55	6836.56	7456

Unit of measurement: Birth rate: number of births per 1000 women of age 15 to 44 CO - mg/m³; other pollutants - µg/m³; Wind speed - km/h; Heating Degree Days - none.

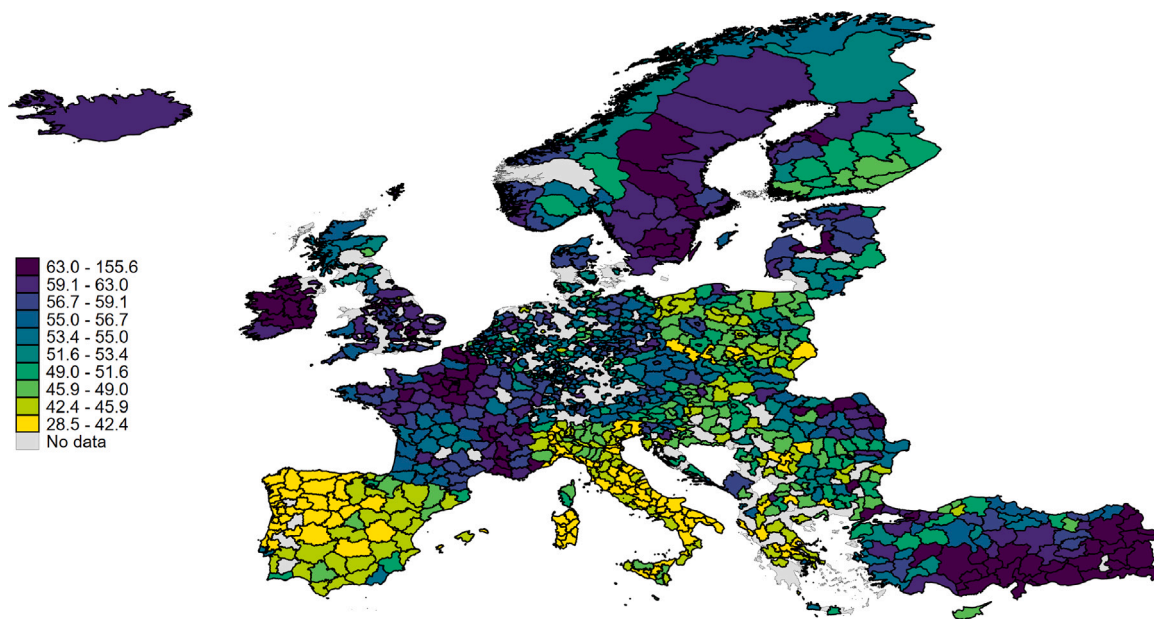


Fig. 1. Average birth rate in regions (2013–2020).
Data source: EUROSTAT. Birth rate: number of births per 1000 women of age 15 to 44.

4. Results

4.1. Descriptive results

First, in [Table 1](#) the descriptive statistics of the main variables are presented. The reported statistics refer to yearly values by regions, except wind speed statistics which refer to the NUTS 2 level.

There is substantial variability in birth rates and pollution concentration levels not only at the country level but also at the regional level. This is shown by the maps in [Figs. 1](#) and [2](#). The maps showing the rest of the pollutants are presented in [Figures A.16](#) to [A.24](#) in the Appendix.

The yearly observations reported in the Appendix ([Table A.11](#)) show that the concentration of NO₂ substantially decreased in the observation period, whereas the concentration of other pollutants such as PM₁₀ and O₃ remained more or less unchanged.

[Figures A.4](#) and [A.5](#) in the Appendix depict the raw associations between the concentration levels of each pollutant in year *t* and the log birth rates in year *t*+1 at the regional level. The raw associations between birth rates and pollution concentrations show diverse nonlinear patterns, including strong negative, strong positive, roughly zero, and varying associations between birth rates and pollution concentrations. These patterns reflect not only the direct negative effect of pollution on birth rates, but also the various indirect effects of secondary factors such as industrial activity and economic prosperity. We provide further descriptive

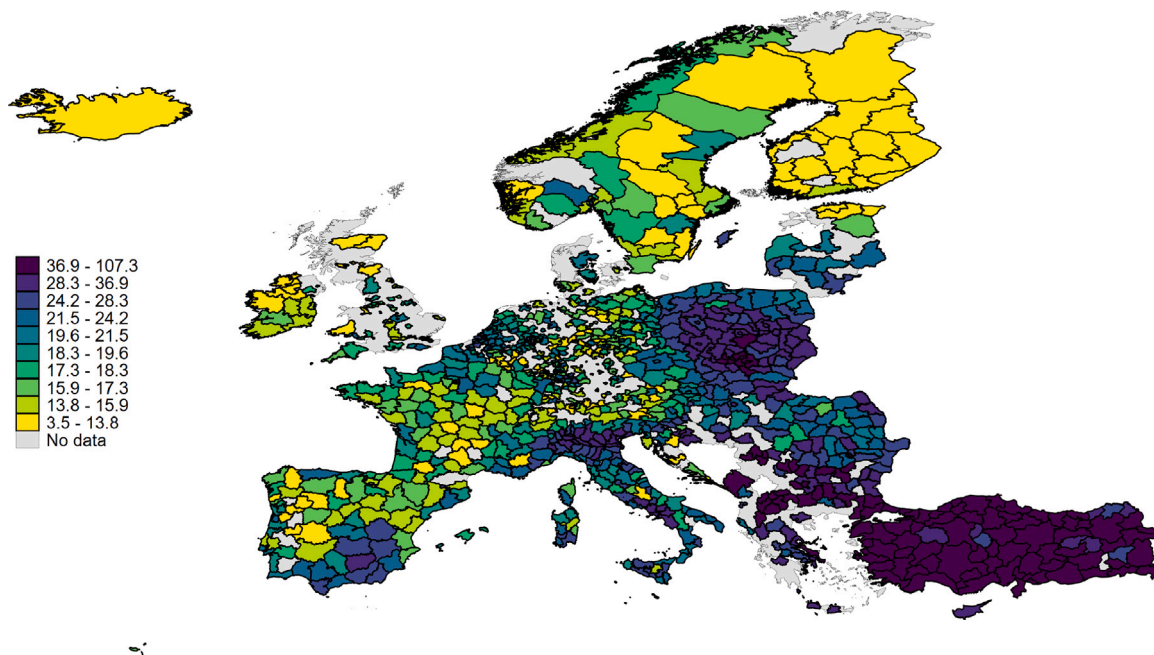


Fig. 2. Average PM₁₀ pollution in regions (2013–2020).
Data source: European Environment Agency. Unit of measurement: µg/m³

figures about pollution concentrations in the Appendix; see Figure A.6 in the Appendix on the associations between instruments and pollution concentration levels and Figures A.7 and A.8 in the Appendix on the pollution distributions relative to the European air quality standards.

4.2. Main results

4.2.1. OLS and 2SLS

The main results from the regression Eqs. (1), (2), (3), and the second stage results of the 2SLS regressions are reported in Table 2. In these regressions, air pollution concentrations are measured as mean concentrations. The pollution variables are standardized to have zero mean and the standard deviation equal to one for easier interpretation of the results. In the first, second, and third columns, the simple OLS regression results are reported. Nevertheless, the OLS estimates are biased, and thus we do not interpret these coefficients.

In Column 4, we present the second-stage results of the 2SLS regression. Here, the point estimate of the PM Factor_{t-1} is larger than the OLS estimate and is significant at the 5% level, whereas the PM Factor_{t-2} is significant at the 1% level. The estimate indicates that if the average PM pollution concentration increases by 1 standard deviation, birth rates decrease by 14.1% the following year and by an additional 17.2% two years later. The 95% confidence intervals are [1.1%; 27%] and [4.1%; 30.3%], respectively. According to our data, a 1 standard deviation increase in this measure is approximately equivalent to a 0.06 to 0.3 mg/m³ increase in CO, a 3 to 12 µg/m³ increase in PM_{2.5}, and a 7 to 13 µg/m³ increase in PM₁₀ concentration levels.⁴ For context, in March 2024, the European Council and Parliament agreed to reduce the annual limit value of PM_{2.5} from 25 to 10 µg/m³ (EC, 2024). The rest of the pollutants have an insignificant effect on fertility in the 2SLS regression.

In the 2SLS regression, we choose a subset of possible instruments so that the included instruments are valid. First, it is important that the included instruments are relevant, and we tested this with the F statistics of each first-stage regression, as well as checking the significance of the included instruments. Table A.13 in the Appendix summarizes the first stage results of the 2SLS regression presented in Table 2. The F-statistics are significant at the 1% level in each first stage regression, except that of SO_{2, t-2} which is significant at 10%. In addition, each instrument is significant at 1% in at least one of the first-stage regressions. Second, since we are using more instruments than the number of endogenous variables, we test whether additional instruments are exogenous. The

⁴ Recall that the PM Factor is a composite index generated using principal factor method for PM_{2.5}, PM₁₀ and CO concentration levels. This is a nonlinear transformation, and, as a result, a 1 SD increase in PM Factor can be associated with various levels of increases in the basic pollutants, depending on what level of concentration we start from.

Table 2
OLS and 2SLS regression estimates of pollution concentrations on log birth rates.

	(1)	(2)	(3)	(4)
L.PM Factor	-0.004** (0.002)	-0.003* (0.002)	-0.003* (0.002)	-0.141** (0.066)
L2.PM Factor		0.000 (0.002)	0.000 (0.002)	-0.172*** (0.067)
L.NO Factor	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	-0.003 (0.051)
L2.NO Factor		0.005*** (0.002)	0.005*** (0.002)	0.023 (0.076)
L.SO ₂	0.002* (0.001)	0.003* (0.001)	0.003* (0.001)	-0.213 (0.201)
L2.SO ₂		0.002 (0.002)	0.002 (0.002)	0.034 (0.239)
Observations	5320	5320	5320	5320
Prob > F	0.093	0.012	0.011	0.067
Clusters	889.000	889.000	889.000	889.000
Model	OLS	OLS	OLS	2SLS
Region FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region linear trend	Yes	Yes	Yes	Yes
GDP	No	No	Yes	Yes

Standard errors in parentheses

Notes: Regressions based on Eq. (1), (2), (3) and 2SLS.

Dependent variable : log birth rate.

Air quality measure: number of days when the pollution concentrations exceeded 125% of the European air quality standards, standardized

L.: first lagged values; L2.: second lagged values.

PM Factor: PM₁₀, PM₂, CO;

NO Factor: NO₂, NO_x, O₃ (Principal factor method).

Robust standard errors clustered at the regional level.

Instruments: Mean wind speed (WS), WS², Number of heating days (HDD),

HDD², HDD², Days(WS > 5 km/h)

Hansen'J test Chi-sq(17) P-value = 0.3711

The full table of estimates is reported in Table A.14 in the Appendix.

* p < 0.10, ** p < 0.05, *** p < 0.01.

p-value of Hansen's J test of overidentification is 0.3711, thus, we cannot reject the null hypothesis of instrument exogeneity.

Lastly, for comparison, we run regressions that include only one pollutant at a time; otherwise, we run the exact same regressions as in Columns 3 and 4 of Table 2. The results of these regressions are reported in Table A.15 in the Appendix. As a result, almost all pollutants have a significant effect in these regressions. This striking difference demonstrates how measuring the effect of only one pollutant can lead to biased estimations.

4.2.2. Robustness checks

In our first robustness check, we use various measures of air pollution concentrations, and report the 2SLS estimation results in Table A.16 in the Appendix. In Column 1 we repeat the results reported in Column 4 of Table 2 for comparison. Column 2 shows the point estimates for the number of days when the pollution concentration exceeded 75% of the concentration limits, and Column 3 shows the results for the number of days when the pollution concentration reached the 10th decile. These results are very similar to the main regression results in magnitude. The effect of PM Factor concentrations are insignificant in Column 2, and significant at the 1% in Column 3. NO Factor and SO₂ have no significant effect on fertility in either specifications.

Second, we check whether using different methods of variable reduction methods in factor variables could affect our results. Table A.17 in the Appendix reports the results for various methods of factorizing. Using the principal-component factor and the iterated principal-factor method, our results remain similar. Whereas, using the maximum likelihood factor method, the NO Factor appears to also be an important pollutant.

Third, we check whether alternative LASSO model specifications lead to different results than the baseline specification. We get very similar results, if we apply different methods to select the λ parameter. Instead of using the minimum of the CV function, we use the "one-standard-error rule" (Hastie et al., 2015); the minimum value of the BIC function; the minimum of the BIC function where models are fit for all lambdas in the grid until the tolerance value is reached; adaptive ridge (adaptive lasso, using the ridge estimator to construct the initial weights in the first lasso); adaptive steps (adaptive lasso with 100 lassos); and adaptive power 1.5 (adaptive lasso, where weights are raised to the 1.5th power). All 7 of these specifications are combined with four air pollution measures, which gives altogether 28 LASSO models. Among the three most important variables, PM₁₀ is included in 14, PM_{2.5} is included in 13, NO_x is included in 11, and O₃ is included in 10 of the 28 models. NO, Pb, and C₆H₆ are not included among the 3 most important variables in any of these models (see Tables A.19, A.20 and A.21 in the Appendix). We also run the LASSO models

Table 3
Heterogeneity by PM₁₀ and GDP (2SLS) (regression estimates of pollution concentrations on log birth rates).

	(1) High PM ₁₀	(2) Low PM ₁₀	(3) High GDP	(4) Low GDP
L.PM Factor	-0.006 (0.017)	-0.016 (0.010)	-0.027* (0.014)	0.010 (0.016)
L2.PM Factor	-0.028 (0.017)	-0.018** (0.009)	-0.014 (0.015)	-0.040** (0.019)
L.NO Factor	0.013 (0.015)	0.009 (0.008)	-0.006 (0.012)	0.008 (0.021)
L2.NO Factor	-0.003 (0.016)	0.014 (0.009)	0.015 (0.012)	0.002 (0.021)
L.SO ₂	-0.060* (0.037)	-0.011 (0.018)	-0.027 (0.036)	-0.025 (0.029)
L2.SO ₂	-0.035 (0.037)	0.005 (0.017)	-0.128** (0.061)	-0.031 (0.028)
Observations	3557	1763	2976	2344
Prob > F	0.148	0.048	0.001	0.041
Clusters	594.000	295.000	497.000	392.000
Model	2SLS	2SLS	2SLS	2SLS
Region FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region linear trend	Yes	Yes	Yes	Yes
GDP	Yes	Yes	Yes	Yes

Standard errors in parentheses

Notes: Dependent variable : log birth rate. Pollution measure: mean.

L.: first lagged values; L2.: second lagged values.

PM Factor: PM₁₀, PM₂, CO; NO Factor: NO₂, NO_x, O₃ (Principal factor method).

Robust standard errors clustered at the regional level.

Within high and low GDP groups we include actual levels of GDP as a control variable.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

with other seeds, and the results are very similar. The results are omitted from the article, but are available upon request.

Fourth, we plot the pollution values and birth rates by year, see Figures A.9 to A.15 in the Appendix. In COVID-19 years there is a slight drop in mean pollution values due to travel restrictions, and also a drop in birth rates. In a robustness check, we run the regressions omitting these two years. The results are reported in Table A.18 in the Appendix. The coefficients remain significant, and the magnitudes are similar to our main results.

Fifth, in our main specification, we choose the instruments so that they are valid. In this part, we test whether choosing a different set of instruments would have led to different coefficient estimates. In Table A.22 in the Appendix we present nested models where specification (1) includes each possible instrument, and we randomly drop some of the instruments in each round. The results show that our model is stable and the results are robust to the choice of instruments. Note that, due to the randomness of the model selection, specification 7 is not the same as our main specification.

4.2.3. Heterogeneity analysis

In this section, we present a heterogeneity analysis, with the results reported in Table 3. In the first two regressions (Columns 1 and 2) we included regions of which the mean yearly average PM₁₀ concentrations are above or under the overall median yearly average. The point estimates of PM₁₀ are not statistically different in the two groups, but the results are more significant in the low pollution areas. The other pollutants are insignificant, similar to the main specification (see Column 4 of Table 2).

We divide regions by GDP in a similar fashion (see Figure A.25 in the Appendix). PM₁₀ one year ago is weakly or not significant. PM₁₀ two years ago has a significant negative effect in the low GDP regions, but not in the high GDP regions. This result is probably due to the higher quality of health services or to the generally better health status of the population in high GDP regions. Similarly to our main result, the NO Factor is not significant in either of the two groups. However, the point estimate of SO₂ in the high GDP group suggests a large negative effect and it is highly significant.

4.2.4. LASSO

The results of the baseline LASSO estimations are reported in Table 4. λ is selected by using a cross-validation function, and the grid for λ is set to 10,000. The dependent variable is log birth rate and the explanatory variables are each single air pollutants (in years $t - 1$ and $t - 2$) that we have data of (including NO, Pb, and C₆H₆), year fixed effects, country fixed effects, and country linear trends. Only the results for the air pollutants are reported in the table. Based on the absolute value of the LASSO coefficients, the first, second, and third most important pollutants are indicated in the models. Less important pollutants that were also kept in the model are marked with an x. The coefficients are not reported or interpreted here because these are subject to omitted variables bias such as the OLS regression results. We report results for various measures of air pollution. In Column 1 we report the LASSO results with pollution measures referring to the number of days when the pollution exceeded 125% of the EU air quality standards

Table 4
LASSO results (λ selection with cross-validation).

	AQS 125%	AQS 175%	D10	Mean
PM ₁₀	1	1	1	1
PM _{2.5}	3	3	3	3
CO			x	x
NO ₂	x	x	x	x
NOx	2	2	2	x
O ₃	x		x	2
SO ₂	x	x	x	x
Pb	x	x	x	x
C ₆ H ₆	x	x	x	x
NO	x	x		

Notes: The grid for λ is set to 10,000. Seed: 1234. Dependent variable: log birth rate. Independent variables: air pollutants (in years $t - 1$ and $t - 2$), year FE*, country FE*, country linear trend*. * Omitted from the Table. The rule used to select λ : CV - minimum of the CV function. Measures of air pollution: (1) Number of days concentration exceeds 125% of the Air quality standard, (2) Number of days concentration exceeds 175% of the Air quality standard, (3) Number of days concentration in Decile 10, (4) Mean. Results: 1: included in the model with the highest importance (highest lasso coefficient in absolute value), 2: second, 3: third, x: included in the model with lower importance. The results of other LASSO model specifications are reported in Tables A.19 and A.20 in the Appendix.

(AQS 125%). In Column 2, AQS 175% is reported, which indicates days with very high pollution concentrations. In Column 3, the pollution is measured with the number of days when the concentrations were highest, 10th decile. In Column 4 the means of the pollution concentrations are used.

Among the three most important variables, PM₁₀ and PM_{2.5} are included in four, NOx is included in three, and O₃ is included in one of the four models. CO concentrations are included in only two of the models, indicating that the 2SLS results on the PM Factor are most likely driven by PM₁₀ and PM_{2.5} concentrations.

The LASSO estimates are suitable to judge the importance of the three variables that we omitted from the OLS and 2SLS regressions due to a low number of observations (NO, Pb, and C₆H₆). These variables are either not included or are among the less important variables in each LASSO model. This result suggests that we likely have not excluded any crucial pollution variables from the OLS and 2SLS regressions.

5. Simulation

Finally, we calculate the predicted birth rate differences between the actual data and a hypothetical scenario with improved levels of particle pollution. In the improved pollution scenario we assume that particulate matter pollution uniformly decreases by 5% in each area, whereas the concentrations of NO Factor and SO₂ concentrations remain at the original level. We predict birth rates for each region and each year, based on our main point estimates (Column 4 of Table 2). Fig. 3 shows the resulting differences compared to the actual birth rates. These results show that in most European regions birth rates would likely increase if the regions complied with the stricter regulations adopted in February 2024 by the European Council (European Council, 2024). Eastern European regions and Northern regions of Italy would likely benefit most from reducing PM₁₀ and PM_{2.5} pollution levels in terms of fertility rates.

6. Discussion

In this paper, we investigate the impact of different types of ambient pollutants on birth rates in Europe. Estimating the causal effect presents challenges due to omitted variable bias resulting from unaccounted pollutants and unmeasured factors. We address this by incorporating the ten most important pollutants into our analysis and employing wind speed and heating days as instrumental variables for pollution concentrations. Previous estimates of the influence of air pollution on fertility have generally been limited to limited geographic and temporal scopes. Our study expands upon this by examining 657 regions across Europe and its neighboring areas, drawing from up to six years of data, thereby enhancing the external validity of our findings. Our results remain robust across various specifications and robustness tests.

We find that particulate matter concentrations, specifically PM_{2.5} and PM₁₀ have a significant effect on birth rates. After controlling for these effects, the rest of the pollutants are found to exert an insignificant effect on fertility. The PM Factor coefficient is significant at the 1% level and it suggests that an increase by 1 SD would result in a 14.1% drop in birth rates next year and a 17.2% drop two years later. These results are robust across various specifications. The effects of other pollutants on birth rates are insignificant in most specifications.

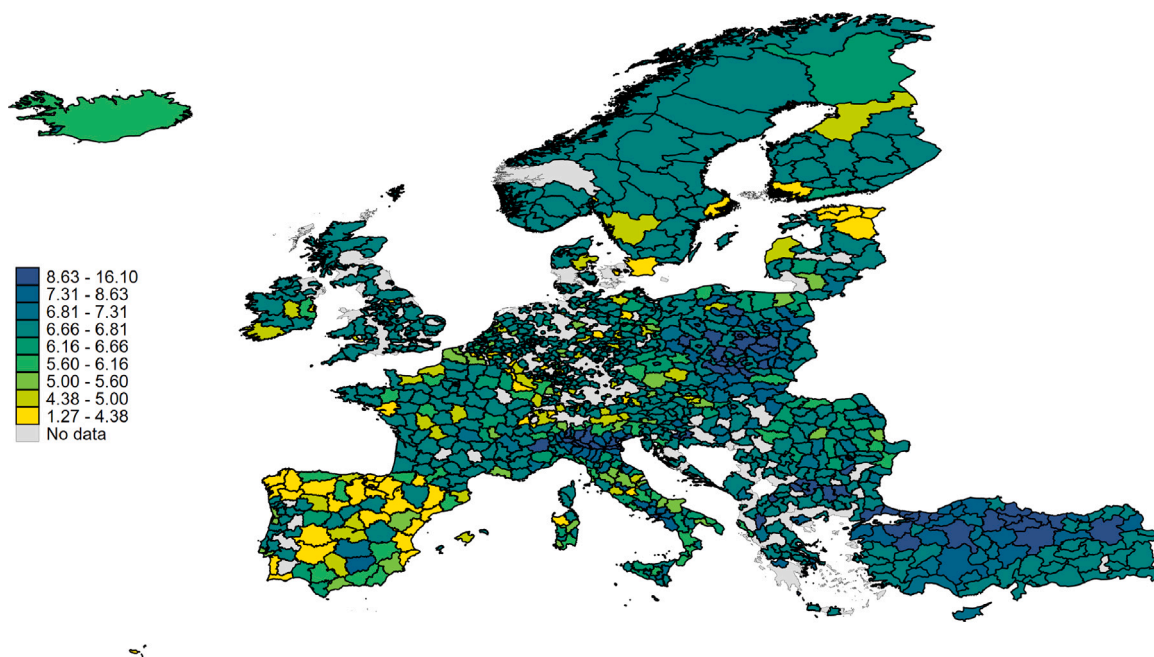


Fig. 3. Predicted birth rate growth (%) as a result of a substantial decrease in particulate matter pollution.
 Notes: Linear prediction of birth rate changes if PM₁₀ decreases uniformly by 5% in each area.

The previous multi-pollutant literature finds contradictory results as to which pollutants matter the most to fertility. These studies find significant associations between fertility and PM_{2.5} (Slama et al., 2013; Faiz et al., 2012), NO₂ (Slama et al., 2013; Faiz et al., 2012; Green et al., 2009), SO₂ (Faiz et al., 2012), CO (Faiz et al., 2012) and O₃ (Green et al., 2009). Our quasi-experimental results support the results related to particulate matter concentrations, but not the other pollutants. These results are similar to the results of Nieuwenhuijsen et al. (2014), who find that PM_{coarse} (PM greater than 2.5 but smaller than 10 µg/m³) drives the fertility effects when NO_x, NO₂, PM_{2.5}, PM₁₀ and PM_{coarse} are jointly included in the regressions.

We add to the previous multi-pollutant literature by going beyond associational relations and carry out a quasi-experimental analysis to uncover causal effects. We also extend the analysis from one region or country and (or) a short time period to the whole European region and 6 years, and show that the results found hold to a multitude of regions and years.

To put our findings into perspective, we make a back-of-the-envelope calculation: in order to achieve a 14.1% increase in fertility throughout the EU, an additional child cash benefit would have to be introduced in the value of about EUR 9.9 billion.⁵ This would be a fertility effect approximately equivalent in magnitude to that induced by the 15 µg/m³ decrease in PM_{2.5} concentration levels prescribed by the agreement of the European Council and Parliament (EC, 2024).

Our findings are significant for governments striving to address air pollution despite resource constraints. The results indicate that there can be substantial synergy between environmental and population policies. Therefore, policymakers can improve the efficiency of expenditure structures by considering these complementary effects.

CRedit authorship contribution statement

Árpád Stump: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation.
Ágnes Szabó-Morvai: Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

⁵ It would cost approximately EUR 18 thousand in additional cash transfers to increase number of births by one (Szabó-Morvai et al., 2019). To increase fertility in the EU by 14.1%, it would take 547 thousand additional births.

Acknowledgments

We are grateful to Bálint Herczeg for the data collection. We thank Anna Adamecz, Bíró Anikó and the participants of the Health and Society Lendület Group Meeting, the ESPE Annual Meeting, the EALE Annual Meeting and the Göttingen Development Economics Seminar for their insightful comments.

Funding

The authors gratefully acknowledge financial support from the Hungarian National Research Development and Innovation Office, Grant no. FK131422, the Lendület programme of the Hungarian Academy of Sciences (LP2018-2/2018). Supported by the University of Debrecen Program for Scientific Publication. The sponsors did not have any role during the design, analysis, and the writing of this study.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jeem.2024.103111>.

References

- Alderotti, Giammarco, Vignoli, Daniele, Baccini, Michela, Matysiak, Anna, 2021. Employment instability and fertility in Europe: A meta-analysis. *Demography* (ISSN: 0070-3370) 58 (3), 871–900. <http://dx.doi.org/10.1215/00703370-9164737>.
- Alexander, Diane, Schwandt, Hannes, 2022. The impact of car pollution on infant and child health: Evidence from emissions cheating. *Rev. Econ. Stud.* (ISSN: 0034-6527) 89 (6), 2872–2910. <http://dx.doi.org/10.1093/restud/rdac007>.
- Angrist, Joshua, Frandsen, Brigham, 2019. *Machine Labor*. Technical report, National Bureau of Economic Research.
- Arceo, Eva, Hanna, Rema, Oliva, Paulina, 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from Mexico city. *Econ. J.* (ISSN: 1468-0297) 126 (591), 257–280. <http://dx.doi.org/10.1111/ecej.12273>, URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/ecej.12273>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecej.12273>.
- Barreca, Alan, Deschenes, Olivier, Guldi, Melanie, 2018. Maybe next month? Temperature shocks and dynamic adjustments in birth rates. *Demography* (ISSN: 1533-7790) 55 (4), 1269–1293. <http://dx.doi.org/10.1007/s13524-018-0690-7>.
- Benmarhnia, Tarik, Bharadwaj, Prashant, Romero, Mauricio, n.d. Using instrumental variables under partial observability of endogenous variables for assessing effects of air pollution on health, URL <https://mauricio-romero.com/pdfs/papers/partialiv.pdf>.
- Böheim, René, Stöllinger, Philipp, 2020. Decomposition of the gender wage gap using the LASSO estimator. *Appl. Econ. Lett.* 1–12.
- Bové, Hannelore, Bongaerts, Eva, Slenders, Eli, Bijmens, Esmée M., Saenen, Nelly D., Gyselaers, Wilfried, Van Eyken, Peter, Plusquin, Michelle, Roef-faers, Maarten B.J., Ameloot, Marcel, Nawrot, Tim S., 2019. Ambient black carbon particles reach the fetal side of human placenta. *Nat. Commun.* (ISSN: 2041-1723) 10 (1), 3866. <http://dx.doi.org/10.1038/s41467-019-11654-3>, URL <https://www.nature.com/articles/s41467-019-11654-3>. Number: 1 Publisher: Nature Publishing Group.
- Buckles, Kasey S., Hungerman, Daniel M., 2013. Season of birth and later outcomes: Old questions, new answers. *Rev. Econ. Stat.* (ISSN: 0034-6535) 95 (3), 711–724. http://dx.doi.org/10.1162/REST_a_00314, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3777829/>.
- Campbell, Jared M., Lane, Michelle, Owens, Julie A., Bakos, Hassan W., 2015. Paternal obesity negatively affects male fertility and assisted reproduction outcomes: a systematic review and meta-analysis. *Reprod. BioMed. Online* (ISSN: 1472-6483) 31 (5), 593–604. <http://dx.doi.org/10.1016/j.rbmo.2015.07.012>, URL <https://www.sciencedirect.com/science/article/pii/S1472648315003776>.
- Chay, Kenneth Y., Greenstone, Michael, 2003. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession*. *Q. J. Econ.* (ISSN: 0033-5533) 118 (3), 1121–1167. <http://dx.doi.org/10.1162/0033553030698513>.
- Conforti, Alessandro, Mascia, Marika, Cioffi, Giuseppina, De Angelis, Cristina, Coppola, Giuseppe, De Rosa, Pasquale, Pivonello, Rosario, Alviggi, Carlo, De Placido, Giuseppe, 2018. Air pollution and female fertility: a systematic review of literature. *Reprod. Biol. Endocrinol.: RB& E* (ISSN: 1477-7827) 16 (1), 117. <http://dx.doi.org/10.1186/s12958-018-0433-z>.
- Currie, Janet, Schwandt, Hannes, 2016. The 9/11 dust cloud and pregnancy outcomes: A reconsideration. *J. Hum. Resour.* (ISSN: 0022-166X) 51 (4), 805–831. <http://dx.doi.org/10.3368/jhr.51.4.0714-6533R>, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5421999/>.
- Currie, Janet, Walker, Reed, 2011. Traffic congestion and infant health: Evidence from E-ZPass. *Am. Econ. J.: Appl. Econ.* (ISSN: 1945-7782) 3 (1), 65–90, URL <https://www.jstor.org/stable/25760246>. Publisher: American Economic Association.
- Deryugina, Tatyana, Heutel, Garth, Miller, Nolan H., Molitor, David, Reif, Julian, 2019. The mortality and medical costs of air pollution: Evidence from changes in wind direction. *Amer. Econ. Rev.* (ISSN: 0002-8282) 109 (12), 4178–4219. <http://dx.doi.org/10.1257/aer.20180279>, URL <https://www.aeaweb.org/articles?id=10.1257/aer.20180279>.
- Doepke, Matthias, Hannusch, Anne, Kindermann, Fabian, Tertilt, Michèle, 2023. The economics of fertility: a new era. In: Lundberg, Shelly, Voena, Alessandra (Eds.), *Handbook of the Economics of the Family*. In: *Handbook of the Economics of the Family*, Volume 1, vol. 1, North-Holland, pp. 151–254. <http://dx.doi.org/10.1016/bs.hefam.2023.01.003>, URL <https://www.sciencedirect.com/science/article/pii/S2949835X23000034>.
- EC, 2024. Air quality: Council and Parliament strike deal to strengthen standards in the EU. *Consilium* URL <https://www.consilium.europa.eu/en/press/press-releases/2024/02/20/air-quality-council-and-parliament-strike-deal-to-strengthen-standards-in-the-eu/>.
- ECA, 2018. Air Pollution: Our Health Still Insufficiently Protected. *Special Report 23/2018*, European Court of Auditors, URL <https://www.eca.europa.eu/en/Pages/DocItem.aspx?did=46723>.
- European Commission, 2020. Copernicus climate change service. URL <https://climate.copernicus.eu/>.
- EuropeanCouncil, 2024. Air quality. URL <https://www.consilium.europa.eu/en/policies/air-quality/>.
- EUROSTAT, 2021. Heating and cooling degree days - statistics. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Heating_and_cooling_degree_days_statistics#Heating_and_cooling_degree_days_by_NUTS_3_regions.
- EUROSTAT, 2022. Fertility (National level) (demo_fer). URL <https://ec.europa.eu/eurostat/web/main/data/database>.
- Faiz, Ambarina S., Rhoads, George G., Demissie, Kitaw, Kruse, Lakota, Lin, Yong, Rich, David Q., 2012. Ambient air pollution and the risk of stillbirth. *Am. J. Epidemiol.* (ISSN: 1476-6256) 176 (4), 308–316. <http://dx.doi.org/10.1093/aje/kws029>.
- Fluchtmann, Jonas, Glenny, Anita M, Harmon, Nikolaj, Maibom, Jonas, 2020. *The Gender Application Gap: Do Men and Women Apply for the Same Jobs?*. Technical report, Mimeo University of Copenhagen.

- Fowler, David, Brimblecombe, Peter, Burrows, John, Heal, Mathew R., Grennfelt, Peringe, Stevenson, David S., Jowett, Alan, Nemitz, Eiko, Coyle, Mhairi, Liu, Xuejun, Chang, Yunhua, Fuller, Gary W., Sutton, Mark A., Klimont, Zbigniew, Unsworth, Mike H., Vieno, Massimo, 2020. A chronology of global air quality. *Phil. Trans. R. Soc. A* 378 (2183), 20190314. <http://dx.doi.org/10.1098/rsta.2019.0314>, URL <https://royalsocietypublishing.org/doi/10.1098/rsta.2019.0314>. Publisher: Royal Society.
- Frutos, Víctor, González-Comadrán, Mireia, Solà, Ivan, Jacquemin, Benedicte, Carreras, Ramón, Vizcaíno, Miguel A. Checa, 2015. Impact of air pollution on fertility: a systematic review. *Gynecol. Endocrinol.* (ISSN: 0951-3590) 31 (1), 7–13. <http://dx.doi.org/10.3109/09513590.2014.958992>.
- Godzinski, Alexandre, Suarez Castillo, Milena, 2021. Disentangling the effects of air pollutants with many instruments. *J. Environ. Econ. Manag.* (ISSN: 0095-0696) 109, 102489. <http://dx.doi.org/10.1016/j.jeem.2021.102489>, URL <https://www.sciencedirect.com/science/article/pii/S0095069621000668>.
- Green, Rochelle S., Malig, Brian, Windham, Gayle C., Fenster, Laura, Ostro, Bart, Swan, Shanna, 2009. Residential exposure to traffic and spontaneous abortion. *Environ. Health Perspect.* (ISSN: 0091-6765) 117 (12), 1939–1944. <http://dx.doi.org/10.1289/ehp.0900943>, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2799470/>.
- Gutierrez, Emilio, 2015. Air quality and infant mortality in Mexico: evidence from variation in pollution concentrations caused by the usage of small-scale power plants. *J. Popul. Econ.* (ISSN: 1432-1475) 28 (4), 1181–1207. <http://dx.doi.org/10.1007/s00148-015-0539-y>.
- Hajdu, Tamás, Hajdu, Gábor, 2021. Temperature, climate change, and birth weight: evidence from Hungary. *Popul. Environ.* (ISSN: 1573-7810) 43 (2), 131–148. <http://dx.doi.org/10.1007/s11111-021-00380-y>.
- Hastie, T., Tibshirani, R., Wainwright, M., 2015. *Statistical Learning with Sparsity: The Lasso and Generalizations*. CRC Press, Boca Raton, FL.
- Jahnke, Johanna R., Messier, Kyle P., Lowe, Melissa, Jukic, Anne Marie, 2022. Ambient air pollution exposure assessments in fertility studies: a systematic review and guide for reproductive epidemiologists. *Curr. Epidemiol. Rep.* (ISSN: 2196-2995) 9 (2), 87–107. <http://dx.doi.org/10.1007/s40471-022-00290-z>.
- James, Gareth, Witten, Daniela, Hastie, Trevor, Tibshirani, Robert, 2013. *An Introduction to Statistical Learning*, vol. 112, Springer.
- Knittel, Christopher R., Miller, Douglas L., Sanders, Nicholas J., 2016. Caution, drivers! children present: Traffic, pollution, and infant health. *Mass. Inst. Technol. Press* (ISSN: 0034-6535) URL <https://dspace.mit.edu/handle/1721.1/113913>. Accepted: 2018-03-02T19:11:39Z Publisher: MIT Press.
- Lam, D.A., Miron, J.A., 1996. The effects of temperature on human fertility. *Demography* (ISSN: 0070-3370) 33 (3), 291–305.
- Legro, Richard S., Sauer, Mark V., Mottla, Gilbert L., Richter, Kevin S., Li, Xian, Dodson, William C., Liao, Duaping, 2010. Effect of air quality on assisted human reproduction. *Hum. Reprod.* (Oxf. England) (ISSN: 1460-2350) 25 (5), 1317–1324. <http://dx.doi.org/10.1093/humrep/deq021>.
- Levine, Hagai, Jørgensen, Niels, Martino-Andrade, Anderson, Mendiola, Jaime, Weksler-Derri, Dan, Mindlis, Irina, Pinotti, Rachel, Swan, Shanna H., 2017. Temporal trends in sperm count: a systematic review and meta-regression analysis. *Hum. Reprod. Update* (ISSN: 1460-2369) 23 (6), 646–659. <http://dx.doi.org/10.1093/humupd/dmx022>.
- Lleras-Muney, Adriana, 2010. The needs of the army: Using compulsory relocation in the military to estimate the effect of air pollutants on children's health. *J. Hum. Resour.* 45 (3), 549–590. <http://dx.doi.org/10.3368/jhr.45.3.549>, (ISSN: 0022-166X, 1548-8004) URL <https://jhr.uwpress.org/content/45/3/549>. Publisher: University of Wisconsin Press Section: Articles.
- Nieuwenhuijsen, Mark J., Basagaña, Xavier, Davdand, Payam, Martinez, David, Cirach, Marta, Beelen, Rob, Jacquemin, Bénédicte, 2014. Air pollution and human fertility rates. *Environ. Int.* (ISSN: 0160-4120) 70, 9–14. <http://dx.doi.org/10.1016/j.envint.2014.05.005>, URL <https://www.sciencedirect.com/science/article/pii/S0160412014001470>.
- Nobles, Carrie J, Schisterman, Enrique F, Ha, Sandie, Buck Louis, Germaine M, Sherman, Seth, Mendola, Pauline, 2018. Time-varying cycle average and daily variation in ambient air pollution and fecundability. *Hum. Reprod.* (ISSN: 0268-1161) 33 (1), 166–176. <http://dx.doi.org/10.1093/humrep/dex341>.
- NUTS, 2024. Nomenclature of territorial units for statistics. Wikipedia URL https://en.wikipedia.org/w/index.php?title=Nomenclature_of_Territorial_Units_for_Statistics&oldid=1234853977. Page Version ID: 1234853977.
- Oldereid, Nan B, Wennerholm, Ulla-Britt, Pinborg, Anja, Loft, Anne, Laivuori, Hannele, Petzold, Max, Romundstad, Liv Bente, Söderström-Anttila, Viveca, Bergh, Christina, 2018. The effect of paternal factors on perinatal and paediatric outcomes: a systematic review and meta-analysis. *Hum. Reprod. Update* (ISSN: 1355-4786) 24 (3), 320–389. <http://dx.doi.org/10.1093/humupd/dmy005>.
- Qian, Hong, Xu, Qiaoqiao, Yan, Wenkai, Fan, Yun, Li, Zhi, Tao, Chengzhe, Zhang, Feng, Lu, Chuncheng, 2022. Association between exposure to ambient air pollution and semen quality in adults: a meta-analysis. *Environ. Sci. Pollut. Res.* (ISSN: 1614-7499) 29 (7), 10792–10801. <http://dx.doi.org/10.1007/s11356-021-16484-9>.
- Schwartz, Joel, Austin, Elena, Bind, Marie-Abele, Zanobetti, Antonella, Koutrakis, Petros, 2015. Estimating causal associations of fine particles with daily deaths in Boston. *Am. J. Epidemiol.* (ISSN: 0002-9262) 182 (7), 644–650. <http://dx.doi.org/10.1093/aje/kwv101>.
- Schwartz, Joel, Bind, Marie-Abele, Koutrakis, Petros, 2017. Estimating causal effects of local air pollution on daily deaths: Effect of low levels. *Environ. Health Perspect.* (ISSN: 1552-9924) 125 (1), 23–29. <http://dx.doi.org/10.1289/EHP232>.
- Siegel, Eva L, Ghassabian, Akhgar, Hipwell, Alison E, Factor-Litvak, Pam, Zhu, Yeyi, Steinthal, Hannah G, Focella, Carolina, Battaglia, Lindsey, Porucznik, Christina A, Collingwood, Scott C, Klein-Fedyshin, Michele, Kahn, Linda G, on behalf of program collaborators for Environmental Influences on Child Health Outcomes, 2023. Indoor and outdoor air pollution and couple fecundability: a systematic review. *Hum. Reprod. Update* (ISSN: 1460-2369) 29 (1), 45–70. <http://dx.doi.org/10.1093/humupd/dmac029>.
- Sillanpää, Markus, Hillamo, Risto, Saarikoski, Sanna, Frey, Anna, Pennanen, Arto, Makkonen, Ulla, Spolnik, Zoya, Van Grieken, René, Braniš, Martin, Brunekreef, Bert, et al., 2006. Chemical composition and mass closure of particulate matter at six urban sites in Europe. *Atmos. Environ.* 40, 212–223.
- Slama, Rémy, Bottagisi, Sébastien, Solansky, Ivo, Lepeule, Johanna, Giorgis-Allemand, Lise, Sram, Radim, 2013. Short-term impact of atmospheric pollution on fecundability. *Epidemiol. (Camb. Mass.)* (ISSN: 1531-5487) 24 (6), 871–879. <http://dx.doi.org/10.1097/EDE.0b013e3182a702c5>.
- Szabó-Morvai, Ágnes, Balás, Gábor, Bördős, Katalin, Herczeg, Bálint, 2019. Evaluation of Family Policy Measures and Their Impact on Fertility. European Commission Hungary Representation, URL https://hungary.representation.ec.europa.eu/system/files/2021-12/hetfa_fertilitymodels_20190913.pdf.
- Tibshirani, Robert, 1996. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* (ISSN: 0035-9246) 58 (1), 267–288, URL <https://www.jstor.org/stable/2346178>. Publisher: [Royal Statistical Society, Wiley].
- WHO, 2019. Ten threats to global health in 2019. URL <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>.
- WorldBank, 2024. Agriculture, forestry, and fishing, value added (% of GDP) - European Union. URL <https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?locations=EU>.
- Zabrocki, Léo, Alari, Anna, Benmarhnia, Tarik, 2022. Improving the design stage of air pollution studies based on wind patterns. *Sci. Rep.* (ISSN: 2045-2322) 12, 7917. <http://dx.doi.org/10.1038/s41598-022-11939-6>, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9106699/>.