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.*
[b] *HUN-REN KRTK KTI, 1097 Budapest, Tóth Kálmán u. 4.*
[c] *University of Debrecen, 4032 Debrecen, Egyetem tér 1, Hungary*

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 5.1% decrease in birth rates the following year and an additional 5.9% 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.

Keywords: Ambient Air Pollution, Fertility, Instrumental Variables, Particulate Matter
JEL: Q53, J13, I14

---

[0] **Corresponding author**: Agnes Szabo-Morvai (ORCID: 0000-0003-3629-6352). HUN-REN KRTK KTI and University of Debrecen, email: szabomorvai.agnes@krtk.hun-ren.hu, Fax: +36-1-224 6700, Phone: +36-30-407-1224, Postal Address: HUN-REN KRTK KTI, 1097 Budapest, Tóth Kálmán u. 4. Árpád Stump email: stumparpad@gmail.com.

Extended abstract

1. Introduction

In the past 70 years, fertility rates have been falling in most developing countries, raising concern 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., 2019). 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. Previous 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.

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, 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.

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. We collect information about NO_2, NO, NOx, O_3, SO_2, $PM_{2.5}$ and PM_{10}, C_6H_6, Pb, CO.

In our main specification, we examine the concentrations relative to the European Air Quality Standards as of 2023 (AQS). For example, AQS $125\%_{\mathrm{rtp}}$ is the number of days when the concentration of the pollutant p exceeded 125% the relevant concentration limit in year t and region r.

Birth rates are based on Eurostat data, 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 regional level yearly heating degree days (HDD) data are provided by the Joint Research Centre's AGRI4CAST Resources Portal. 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. From the daily observations, we calculate the yearly mean wind speed. We use principal factor analysis to combine highly correlated pollutants. As a result, we are left with three pollutant variables in the main regressions: *PM Factor* (PM_{10}, $PM_{2.5}$, and CO), *NO Factor* (NO_2, NOx and O_3), and SO_2.

3. Empirical method

There can be 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) and others use wind direction and speed, and Arceo et al. (2016) use temperature (thermal inversions) to instrument endogenous ambient air pollution concentrations.

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. 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:

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

where subscript j denotes the number of lags, and k is the k$^{\text{th}}$ 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:

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

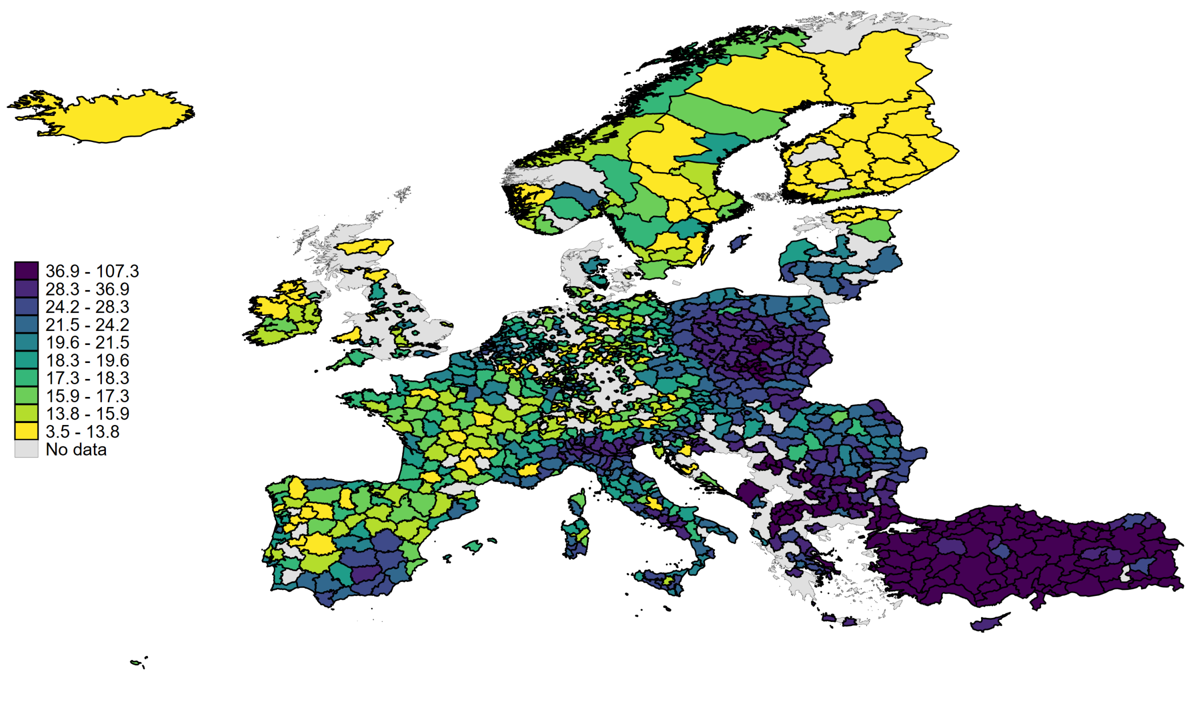
The reduced-form equations are the following:

$$ln(Y_{rt}) = \sum_{j=1}^{2} \sum_{k=1}^{22} (\gamma_{k,t-j} Z_{k,t-j}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \tag{3}$$

We apply various robustness checks and LASSO estimations to show that our results are robust across various specifications.

4. Results

Figure 1: Average PM_{10} pollution in regions (2013-2020)



Data source: European Environment Agency. Unit of measurement: $\mu g/m^3$

We find that particulate matter concentrations, specifically $PM_{2.5}$ and PM_{10} 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 5.1% drop in birth rates the next year and 5.9% two years later. These results are robust across various specifications. The effects of other pollutants on birth rates are insignificant in most specifications.

Table 1: OLS and 2SLS regression estimates of pollution concentrations on log birth rates

	(1)	(2)	(3)	(4)
L.PM Factor	-0.005***	-0.005***	-0.005***	-0.051***
	(0.001)	(0.002)	(0.002)	(0.010)
L2.PM Factor		-0.002	-0.002	-0.059***
		(0.002)	(0.002)	(0.013)
L.SO_2	0.027	0.009	0.009	0.299
	(0.035)	(0.043)	(0.043)	(1.198)
L2.SO_2		-0.052	-0.051	1.683
		(0.038)	(0.038)	(1.686)
L.NO Factor	-0.000	0.000	0.000	-0.005
	(0.001)	(0.001)	(0.001)	(0.011)
L2.NO Factor		0.001	0.001	0.018*
		(0.001)	(0.001)	(0.010)
Observations	5320	5320	5320	5320
Prob > F	0.010	0.000	0.000	0.000
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

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: PM_{10}, PM_2, CO; NO Factor: NO_2, NO_X, O_3 (Principal factor method).

Robust standard errors clustered at the regional level.

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