INSTITUTE OF ECONOMICS
Centre for Economic and Regional Studies

Közgazdaság- és Regionális Tudományi Kutatóközpont
KÖZGAZDASÁG-TUDOMÁNYI INTÉZET

KRTK-KTI WORKING PAPERS | KRTK-KTI MŰHELYTANULMÁNYOK

The Effect of Air Pollution on Fertility Outcomes in Europe

ÁRPÁD STUMP – BÁLINT HERCZEG – ÁGNES SZABÓ-MORVAI

KRTK-KTI WP – 2023/10

May 2023

https://kti.krtk.hu/wp-content/uploads/2023/05/KRTKKTIWP202310.pdf

INSTITUTE OF ECONOMICS, CENTRE FOR ECONOMIC AND REGIONAL STUDIES
EÖTVÖS LORÁND RESEARCH NETWORK (ELKH)
BUDAPEST

ABSTRACT

This paper studies the effect of ambient air pollution on the number of births in the European Union. We collect air pollution data with web scraping technique and utilize variations in wind, temperature, number of heating, and cooling days as instrumental variables. There are 657 NUTS 3 regions included in the regressions, each with 2 to 6 years of observations between 2015 and 2020. Our results show that an increase in the levels of $PM_{2.5}$ - PM_{10} pollution concentration by 1 μg/m^3 (appr. 5-10%) would result in a 9% drop in the number of births next year. CO pollution levels also have a significant although smaller effect. If CO pollution concentration increases by 1 mg/m^3 (appr. 15%) the number of births next year will fall by about 1%. In the heterogeneity analysis, we find that air pollution is more harmful to fertility in countries with already high pollution levels and lower GDP. This latter suggests that healthcare spending and the general level of living standard could be factors that moderate the negative consequences of ambient air pollution. To our knowledge, this is the first article to study the fertility effects of air pollution using an extended number of countries and years and at the same time including more than one air pollutant. As a result, our results have strong external validity. A remarkable novelty of our study compared to the previous literature is that after taking into account the effect of $PM_{2.5}$ - PM_{10} and CO, the rest of the pollutants have much less role in shaping fertility outcomes compared to the findings of the previous literature. This difference is a result of the new method of this study, which examines the pollutants simultaneously instead of examining only one or a few at a time. This result can be important for environmental policies, where the limited resources should target pollution types that have the most detrimental effect on human fertility and health.

JEL codes: Q53, J13, I14
Keywords: ambient air pollution, fertility, instrumental variables

Stump Árpád
Budapesti Corvinus Egyetem
stumparpad@outlook.hu

Szabó-Morvai Ágnes
KRTK KTI and Debreceni Egyetem
szabomorvai.agnes@krtk.hu

Herczeg Bálint
HÉTFA Kutatóintézet
herczegbalint@hetfa.hu

A légszennyezési hatása a születésszámra Európa országaiban

STUMP ÁRPÁD – HERCZEG BÁLINT – SZABÓ-MORVAI ÁGNES

ÖSSZEFOGLALÓ

Ez a tanulmány a környezeti levegő szennyezettségének hatását vizsgálja a születések számára az Európai Unióban. A légszennyezettségi adatokat webscraping technikával gyűjtjük, és instrumentális változóként a szél, a hőmérséklet, a fűtési és hűtési napok számának változásait használjuk. A regressziókban 657 NUTS 3 régió szerepel, mindegyik 2-6 évnyi megfigyeléssel 2015 és 2020 között. Eredményeink azt mutatják, hogy a $PM_{2,5}$ - PM_{10} szennyezettségi koncentráció szintjének 1 μg/m^3-rel (kb. 5-10%) történő emelkedése a születések számának 9%-os csökkenését eredményezi a következő évben. A CO szennyezettségi szinteknek is jelentős, bár kisebb hatása van. Ha a CO-szennyezettség koncentrációja 1 mg/m^3-rel (kb. 15%-kal) nő, a születések száma jövőre mintegy 1%-kal csökken. A heterogenitáselemzés során azt találjuk, hogy a levegőszennyezés károsabb a termékenységre azokban az országokban, ahol már egyébként is magas a szennyezettség és alacsonyabb a GDP. Ez utóbbi arra utal, hogy az egészségügyi kiadások és az általános életszínvonal olyan tényezők lehetnek, amelyek mérsékelhetik a környezeti levegő szennyezettségének negatív következményeit. Tudomásunk szerint ez az első olyan cikk, amely a levegőszennyezés termékenységi hatásait vizsgálja nagy számú ország és év felhasználásával, ugyanakkor egynél több légszennyező anyag bevonásával. Ennek eredményeként eredményeinknek erős a külső érvényessége. Vizsgálatunk egyik figyelemre méltó újdonsága a korábbi irodalomhoz képest, hogy a $PM_{2.5}$ - PM_{10} és a CO hatását figyelembe véve a többi szennyező anyagnak sokkal kisebb szerepe van a termékenység alakulásában, mint ahogyan a korábbi szakirodalomban találják. Ez a különbség a jelen tanulmány új módszerének eredménye, amely egyszerre vizsgálja a szennyező anyagokat ahelyett, hogy egyszerre csak egyet vagy néhányat vizsgálna. Ez az eredmény fontos lehet a környezetvédelmi politikák számára, ahol a korlátozott erőforrásokkal az emberi termékenységre és egészségre leginkább káros szennyezési típusokat kell célozni.

JEL: Q53, J13, I14
Kulcsszavak: légszennyezés, születésszám, instrumentális változók módszere

The Effect of Air Pollution on Fertility Outcomes in Europe

Árpád Stump[a], Bálint Herczeg[b], Ágnes Szabó-Morvai[a,c,d]

[a] *KRTK KTI, 1097 Budapest, Tóth Kálmán u. 4.*
[b] *HETFA Research Istitute*
[c] *University of Debrecen, 4032 Debrecen, Egyetem tér 1, Hungary*
[d] *CERGE-EI Foundation Teaching Fellow*

Abstract

This paper studies the effect of ambient air pollution on the number of births in the European Union. We collect air pollution data with web scraping technique and utilize variations in wind, temperature, number of heating, and cooling days as instrumental variables. There are 657 NUTS 3 regions included in the regressions, each with 2 to 6 years of observations between 2015 and 2020.

Our results show that an increase in the levels of $PM_{2.5}$ - PM_{10} pollution concentration by 1 $\mu g/m^3$ (appr. 5-10%) would result in a 9% drop in the number of births next year. CO pollution levels also have a significant although smaller effect. If CO pollution concentration increases by 1 mg/m^3 (appr. 15%) the number of births next year will fall by about 1%.

In the heterogeneity analysis, we find that air pollution is more harmful to fertility in countries with already high pollution levels and lower GDP. This latter suggests that healthcare spending and the general level of living standard could be factors that moderate the negative consequences of ambient air pollution.

To our knowledge, this is the first article to study the fertility effects of air pollution using an extended number of countries and years and at the same time including more than one air pollutant. As a result, our results have strong external validity.

A remarkable novelty of our study compared to the previous literature is that after taking into account the effect of $PM_{2.5}$ - PM_{10} and CO, the rest of the pollutants have much less role in shaping fertility outcomes compared to the findings of the previous literature. This difference is a result of the new method of this study, which examines the pollutants simultaneously instead of examining only one or a few at a time. This result can be important for environmental policies, where the limited resources should target pollution types that have the most detrimental effect on human fertility and health.

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

---

[0] We thank Anna Adamecz and the participants of the Health and Society Lendület Group meeting for their insightful comments. The authors gratefully acknowledge financial support from the Hungarian National Scientific Research Program (OTKA), Grant no. FK131422 and the Lendület programme of the Hungarian Academy of Sciences (LP2018-2/2018).

1. Introduction

In the past decades, fertility rates have been decreasing all over the world, causing concerns for the sustainability of pension and healthcare systems in many developing countries. At the same time, air pollution is a main environmental concern all around the world. In this research, we examine how much air pollution affects fertility rates. The results of the analysis are of direct policy relevance to the developed countries aiming to combat low fertility.

The 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), lost years of healthy life from ambient air pollution is 0.75 per hundred inhabitants in Europe on average, with some European countries reaching higher levels than China or India (1.7 and 1.6). These pollutants are of high relevance to fertility, as shown by several studies summarized by the meta-study of Frutos et al. (2015). This meta-analysis points out the lack of direct evidence on the effect of air pollution on human live births as a main limitation of the past research. Our study aims to fill this gap by analyzing yearly fertility and air pollution data of 36 European countries at the NUTS 3 region level.

There is a large body of literature studying how air pollution affects fertility in the short and long run. Levine et al. (2017) documents a dramatic, more than 50% decrease in sperm count between 1973-2011 worldwide, and there is direct evidence on the causal relationship between air pollution and declining semen quality such as concentration, count, and motility (Qian et al., 2022). Additionally, according to recent evidence, particle matters from the air may be reaching the placenta which may increase risks to the fetus (Bové et al., 2019). Moreover, Mohallem et al. (2005) confirm in a randomized experiment that the number of pregnancies due to NO_2 and PM_{10} pollution decrease only slightly, but the number of live births fall drastically, by about 33%, and many more studies come to similar conclusions (Faiz et al., 2012; Legro et al., 2010; Mohorovic, Petrovic, Haller and Micovic, 2010; Slama et al., 2013) Conforti et al. (2018) review the available literature on the effect of air pollution on human fertility, but evidence is available only for very short-term effects, such as conception rate after spontaneous intercourse.

Table 1: Effects of air pollutants on fertility according to Conforti et al. Conforti et al. (2018) Table 2.

Type of Pollutant	Population	Effect
NO_2	IVF	Lower live birth rates
	General population	Higher miscarriage rate
CO	General population	Higher stillbirth in second and third trimester
O_3	IVF	Lower live birth rates
$PM_{2.5}$	IVF	Lower pregnancy rates
	General population	Reduced fecundability ratio
PM_{10}	IVF	Higher miscarriage rate
	General population	Higher miscarriage rate
$PM_{2.5\text{-}10}$	General population	Higher miscarriage rate
SO_2	IVF	No effect
	General population	Higher early miscarriage and third trimester still births. Reduced conception rate
Traffic pollutants	General population	Higher miscarriage rate; Higher infertility rates.
Coal combustion products	General population	Higher trend of miscarriage

2. Data

In this research project, we collect air quality data from the European Environment Agency (EEA) using a web scraping technique. The member states upload the air quality data collected by a representative sample of measuring stations. We collected information about nitrogen dioxide (NO_2), nitrogen monoxide (NO), different nitrogen oxides (NO_X), ozone (O_3), sulfur dioxide (SO_2), different sizes of particulate matters ($PM_{2.5}$ and PM_{10}), benzene (C_6H_6, lead (Pb) and carbon monoxide (CO). We downloaded more than 1.1 billion data points (Table A.13). The CO pollution is measured in mg/m^3 and every other pollutant is measured in $µg/m^3$.

We clean the database as follows. We delete observations not on hourly or daily frequency, observations with negative concentration values, and all non-validated observations (mostly missing values). Then we calculate daily averages from the hourly data. We connect the stations to NUTS3 regions using the coordinates of the measurement station. As the countries report only a representative selection of their air quality data to the EEA, we do not have stations in each NUTS3 region. Next, we calculate the average daily concentration across stations for each NUTS3 region. Finally, we calculate the yearly average for each NUTS3 region.

For each type of pollutant, we calculate deciles of the daily pollution levels across every NUTS3 region for the whole observation period. Then we count the number of days in each year and NUTS3 region when the daily average pollution concentration was in the given decile. For instance, Days D10 of $O_{3,rt}$ shows the number of days when the O_3 pollution was in the 10^{th} pollution decile in year t and region r.

We also divide the pollution levels into categories based on the concentration limits (CL) of EU air quality standards (Table A.14). For some pollutants, such as NO and NOX, no daily mean pollution threshold values are set by the EU. For these pollutants, we use the yearly target value or the maximum daily 8-hour mean value. We create 8 categories based on pollution target values [$0\text{-}0.25\times CL$; $0.25\text{-}0.5\times CL$; $0.5\text{-}0.75\times CL$; ... ; $1.75\times CL - \infty$]. E.g., the limit value for $PM_{2.5}$ is 20 $µg/m^3$, and the variable Days I5 of $PM_{2.5,\,rt}$ is the number of days when the pollution level was between 20 $µg/m^3$ and $1.25\cdot 20=25$ $µg/m^3$ in year t and region i.

We add NUTS3-level live birth and female population variables from Eurostat. We calculate the birth rate as the ratio of number of live births and female population on 1^{st} January, between ages 15 and 44.

We use NUTS3 level GDP per person as a control variable.

We use temperature, wind speed, heating degree days (HDD) and cooling degree days

(CDD) to instrument the pollution variables[1]. We got NUTS2 level daily temperature (measured in °C) and wind speed (measured in km/h) data from the Copernicus Climate Change Service. From the daily observations, we calculate the yearly mean temperature and the yearly average wind speed.

The NUTS3 level yearly heating degree days (HDD, if the daily mean temperature (T) is below 15°C then HDD=18-T for that day, 0 otherwise) and cooling degree days (CDD, if the daily mean temperature (T) is above 24°C then CDD=T-21 for that day, 0 otherwise) are available on Eurostat.

Table 2 and 3 show the overall coverage of the variables. Note that we include only the NUTS3 regions that have at least one pollution data observation. HDD and CDD have good coverage, as there is data available for every NUTS3 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. This is not much of an issue, as our empirical model uses pollution level in year $t-1$ to explain birth rate levels in year t.

To increase the number of years and regions included in the final dataset, we impute the missing pollution observations. The imputation strategy is as follows. If there is no observation for a NUTS3 region, we use the NUTS2 average instead. If the NUTS2 level average is also missing, we use NUTS1 level average. If the NUTS1 level average is missing, we impute the country level average. This strategy brings in a measurement error, causing an attenuation bias of the point estimates (Table 4).

---

[1] "Heating degree day (HDD) index is a weather-based technical index designed to describe the need for the heating energy requirements of buildings. Cooling degree day (CDD) index is a weather-based technical index designed to describe the need for the cooling (air-conditioning) requirements of buildings. HDD and CDD are derived from meteorological observations of air temperature, interpolated to regular grids at 25 km resolution for Europe. Calculated gridded HDD and CDD are aggregated and subsequently presented on NUTS-2 level, for 2017 and 2018 also on NUTS-3 level. The severity of the cold in a specific time period taking into consideration outdoor temperature and average room temperature (in other words the need for heating). The calculation of HDD relies on the base temperature, defined as the lowest daily mean air temperature not leading to indoor heating. The value of the base temperature depends in principle on several factors associated with the building and the surrounding environment. By using a general climatological approach, the base temperature is set to a constant value of 15°C in the HDD calculation." See more on these measures on the EUROSTAT webpage.

Table 2: Number of pollution observations by country

country	PM10	SO2	O3	NO2	NOX	CO	C6H6	NO	Pb	PM2.5
AD	7	4	7	7	7	7	0	7	0	0
AL	20	28	31	31	31	32	28	0	0	21
AT	264	213	270	264	79	113	37	254	10	159
BE	210	162	220	239	28	97	166	0	114	212
BG	176	144	109	120	7	93	86	120	52	62
CH	68	42	68	68	68	42	17	68	0	42
CY	8	8	8	8	8	8	8	0	8	8
CZ	112	106	112	112	110	80	9	112	94	112
DE	1685	797	1475	1683	1545	575	274	1710	418	1075
DK	32	27	47	47	41	34	8	47	15	30
EE	32	39	40	39	40	32	24	0	32	40
EL	90	54	70	71	5	44	28	72	8	46
ES	414	415	418	418	418	347	304	418	304	369
FI	126	60	89	104	97	15	7	57	1	79
FR	742	327	750	732	188	167	78	636	26	615
HR	73	43	82	75	73	28	24	0	14	66
HU	100	96	80	96	88	90	66	0	26	53
IE	58	43	57	50	50	27	17	0	0	48
IS	14	14	2	14	13	1	0	14	0	14
IT	795	582	757	797	391	663	626	223	368	719
LT	48	47	64	55	55	40	17	0	0	29
LU	8	8	8	8	7	8	8	8	2	8
LV	32	32	40	32	6	11	26	10	0	23
ME	8	8	8	8	8	8	0	8	0	1
MK	52	53	52	45	45	53	2	47	0	9
MT	16	16	16	16	7	16	13	16	10	16
NL	201	78	202	213	213	39	25	213	6	138
NO	107	65	70	110	98	11	0	98	0	81
PL	572	525	451	537	533	400	301	134	449	446
PT	142	95	141	142	140	50	24	139	7	99
RO	257	283	304	289	37	279	188	0	164	124
RS	35	67	31	59	57	91	2	57	14	3
SE	152	50	104	112	56	20	16	14	0	84
SI	78	33	64	64	57	24	16	0	36	28
SK	64	63	56	64	64	64	64	0	30	61
TR	477	477	236	261	260	195	0	0	0	183
UK	449	209	552	754	753	56	33	573	17	480
XK	7	6	7	5	5	6	0	3	0	7

Table 3: Number of yearly NUTS3 level average birthrate, wind speed, temperature, HDD and CDD onby country

country	birth rate	Temp.	Wind	HDD	CDD
AD	0	0	0	0	0
AL	35	40	40	0	0
AT	245	280	280	280	280
BE	232	264	264	272	272
BG	154	176	176	176	176
CH	70	80	80	0	0
CY	7	8	8	8	8
CZ	98	112	112	112	112
DE	1799	2048	2048	2056	2056
DK	42	48	48	48	48
EE	33	40	40	40	40
EL	140	160	160	160	160
ES	357	416	416	408	408
FI	126	144	144	144	144
FR	686	744	744	744	744
HR	83	104	104	104	104
HU	105	120	120	120	120
IE	56	64	64	64	64
IS	14	16	16	0	0
IT	748	864	864	864	864
LT	56	64	64	64	64
LU	7	8	8	8	8
LV	35	40	40	40	40
ME	7	8	8	0	0
MK	49	56	56	0	0
MT	14	0	0	16	16
NL	203	232	232	232	232
NO	94	128	128	120	120
PL	511	584	584	584	584
PT	140	152	152	144	144
RO	294	336	336	336	336
RS	56	112	112	0	0
SE	147	168	168	168	168
SI	70	80	80	80	80
SK	56	64	64	64	64
TR	560	640	640	0	0
UK	513	800	800	0	0
XK	0	0	0	0	0

Table 4: The number of observation after imputing the missing NUTS3 level pollution observation with the NUTS2, NUTS1 or country level mean pollution.

	NUTS3		**after imputing NUTS2 means**		**after imputing NUTS1 means**		**after imputing country means**	
Variable	**Obs**	**Mean**	**Obs**	**Mean**	**Obs**	**Mean**	**Obs**	**Mean**
PM10	7,731	22.8	8,991	22.4	9,211	22.5	9,225	22.5
SO2	5,319	4.9	8,132	4.2	9,026	4.1	9,211	4.1
O3	7,098	52.0	8,746	51.8	9,018	51.8	9,213	51.6
NO2	7,749	19.6	8,858	19.7	9,085	20.0	9,223	20.1
NOX	5,688	34.0	7,393	32.4	8,295	31.4	9,074	30.8
CO	3,866	4.1	6,731	4.1	7,806	3.6	9,179	3.1
C6H6	2,542	1.2	4,886	1.1	6,563	1.1	8,387	1.1
NO	5,057	11.0	5,952	11.1	6,335	11.2	6,502	11.2
Pb	2,225	0.7	4,024	1.1	5,045	1.6	6,732	3.4
PM2.5	5,590	13.5	8,422	13.5	8,837	13.6	9,124	14.0

In some cases, there is a strong correlation between the yearly pollution levels (Table 5). For example, the correlation coefficient between the $PM_{2.5}$ pollution level and the PM_{10} pollution level is 0.8. NO_2 has a very strong correlation with NOX, the correlation coefficient is 0.9.

Table 5: Correlation matrix of the pollution variables.

	CO	NO	NO2	NOX	O3	PM2	PM10	SO2
CO	1							
NO	-0.0037	1						
NO2	0.0084	0.0002	1					
NOX	0.0039	-0.02	0.9015	1				
O3	0.0261	-0.06	-0.5962	-0.577	1			
PM2	0.0353	0.1922	0.2743	0.2209	-0.327	1		
PM10	0.0474	0.2314	0.2002	0.1712	-0.176	0.806	1	
SO2	0.0649	-0.007	-0.0485	-0.052	0.0803	0.2696	0.3569	1

3. Empirical method

3.1. Baseline specifications

We would like to estimate the effect of air pollution on the number of births. First, we estimate naive regressions of pollution indicators in the previous year on the natural logarithm of the birth rate. It is important that the pollutants are included in the same model because they are correlated and many of them may affect fertility. If one would examine one at a time, the estimates would suffer from omitted variable bias.

Thus, we include mean values of all the pollutants available in the data, except those with a very low number of observations (C_6H_6 and Pb). Some of the pollutants are highly correlated, thus we combined them with factor analysis to avoid the multicollinearity problem. As a result, we have five pollutants: PM factor (including PM_{10} and $PM_{2.5}$), NO_2 factor (including NO_2, NOX and O_3), SO_2, NO and CO. For the details of the factorization, see Appendix.

The observations are aggregated to the year (t) and NUTS 3 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 are unchanged in time, such as social norms that influence environmental consciousness and fertility decisions. 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 NUTS 3 level.

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

We calculate robust standard errors clustered at the NUTS 3 region level (see Abadie et al., 2023).

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

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

The naïve estimation strategy outlined above results in biased point estimates because of omitted factors. Even after including two-way fixed effects and region-specific time trends, there can be time-varying, region-specific factors that correlate with pollution and fertility. To show this, we include NUTS 3 level GDP per inhabitant as a control variable measured with the 2020 purchasing power standard using EUROSTAT data.

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

However, there are 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 n our point estimates.

To circumvent this source of bias, we follow an instrumental variables design. Our instruments are temperature, wind speed, number of heating days, and number of cooling days. These variables have been used as instruments for pollution in the literature before. Knittel, Miller and Sanders (2016) use local weather conditions, Schwartz et al. (2015), Schwartz, Bind and Koutrakis (2017) and Deryugina et al. (2019) use wind direction and speed, and Arceo, Hanna and Oliva (2016) use temperature (thermal inversions) to instrument endogenous ambient air pollution concentrations.

In the previous studies, only one or just a few pollutants were included which only required one or just a few instruments. According to the results of Benmarhnia, Bharadwaj and Romero (n.d.), in this case, using an instrumental variable design does not provide unbiased point estimates, because the exclusion restriction likely does not hold, when the instrument affects the pollutants omitted from the regressions.

In this study, we include 5 pollutant types, thus we need at least 5 instruments. We use non-linear combinations of the four instruments, including their squared and the interactions. We include lagged values of the instruments similarly to the pollutants. In our main specification, one and 2 year lags of the pollutants and the instruments are included.

We run two-stage least squares regressions (2SLS). The first-stage regressions show how strong and significant a relationship the instruments have with pollution concentrations. The first stage for the pollution concentrations one year before birth:

$$P^i_{r,t-1} = \sum_{j=1}^{2} \sum_{k=1,k\neq m}^{4} (\pi_{1k} Z_{k,t-j} + \pi_{2k} Z^2_{k,t-j} + \pi_{3k} Z_{k,t-j} \times Z_{mt}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \quad (4)$$

and the first stage for the pollutants two years before birth:

$$P^i_{r,t-2} = \sum_{j=1}^{2} \sum_{k=1,k\neq m}^{4} (\pi_{1k} Z_{k,t-j} + \pi_{2k} Z^2_{k,t-j} + \pi_{3k} Z_{k,t-j} \times Z_{mt}) + \tau GDP_{rt} + \eta_t + \lambda_r + \lambda_r \times t + \varepsilon_{rt} \quad (5)$$

The intuition is that these weather-related factors affect ambient air pollution concentration and composition. Higher wind speed helps to dissipate high concentrations of ambient air pollution. Higher temperatures can increase ambient air pollution in various ways. First, in heatwaves, wildfires are more frequent which produce high quantities of particles (PM_{10} and $PM_{2.5}$). Second, sunlight and heat induce chemical reactions between primary air pollutants such as nitrogen oxides and oxygen, forming secondary pollutants such as ozone. Also, heat transforms larger particles into smaller and more toxic ones. Third, heat waves come with high atmospheric pressure which keeps air pollution at the ground level and thus increases its concentration.

Lastly, on very hot days air conditioning is more heavily used in buildings and cars which increases car and power plant emissions. Also, in the cold days of winter, the emissions also increase as a result of the heating activity. The cooling and heating activity at a certain temperature may differ by region, which we capture with the number of heating days and the number of cooling days instruments.

The instrumental variables strategy provides unbiased estimates if the exclusion restriction holds. This ultimately consists of two parts. First, these variables are 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. Second, it is important that these weather conditions affect fertility only directly through air pollution and no other channels. In the previous literature, we know of no evidence that wind speed would affect fertility rates. Higher than 25-degree temperature has been shown to affect conception rates negatively (Hajdu and Hajdu, 2022), but the effects are only temporary and result in the alteration of the timing of the pregnancies but do not affect yearly fertility rates. The heating and cooling activities that aim to provide normal ambient temperature conditions to individuals have also not been shown to affect fertility rates.

The reduced-form equations are the following:

$$ln(Y_{rt}) = \sum_{k=1,k\neq m}^{4} (\pi_{1k} Z_{kt} + \pi_{2k} Z_{kt}^2 + \pi_{3k} Z_{kt} \times Z_{mt}) + \tau GDP_{rt} + \eta_t + \delta_r + \lambda_r \times t + \varepsilon_{rt} \tag{6}$$

3.2. Robustness checks

As a first robustness check, we include lead values of pollutants in the regressions as placebo treatments. These should not be significant, as the future values of pollution should not have an effect on fertility rates. Also, the inclusion of these future values should not alter our main estimation results.

Second, we include other measures of ambient air pollution concentrations. Instead of the mean of the pollutants, we use the maximum pollution and the number of days in the highest deciles of pollution as described in Section 2.

Third, we check whether including various deciles and intervals of the pollution concentrations show a reasonable pattern of effects. We expect that the highest pollution concentration deciles should have a higher effect on fertility.

3.3. Heterogeneity

In our unique data, we have many European Union regions included, thus we are able to present a heterogeneity analysis. We divide the sample by the average levels of PM concentrations through the observation period. The high pollution subsample includes NUTS 3 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. Still, about 30% of the regions are in the "high pollution - high GDP" or the "low pollution - low GDP" categories.

4. Results

4.1. Descriptive results

First, we report some descriptive results of our dataset. In Table 6 the descriptive statistics of the main variables are presented. The reported values are yearly average values by NUTS3 region. Note that wind speed and temperature variables are NUTS2-level data. NUTS3 regions in the same NUTS2 area have the same wind speed and temperature.

The first map (Figure 1) shows the dispersion of birth rates across the NUTS3 regions of the EU. There is substantial variability in the outcome variable not only at the country but also at the regional level. The average regional PM_{10} pollution concentrations are shown in Figure 2.

Table 7 reports the yearly average pollution levels over the regions for each pollutant. The concentration of Pb have substantially decreased in the observation period, whereas other pollutants such as PM_{10} and O_3 remained unchanged.

4.2. Main results

The results of the regressions in Equations 1, 2, 3, and the second stage of the 2SLS regressions are reported in Table 8. In the first, second, and third models, the simple OLS regression results are reported and suggest a moderate and significant effect of the particulate matter concentrations and a very slight effect of CO concentration.

In Model 4, we report the results of the 2SLS regression. Here the point estimate of the particular matter is larger compared to the OLS estimate. The PM coefficient is significant at the 1% level, and it suggests that an increase in the levels of $PM_{2.5}$ - PM_{10} pollution concentration by 1 (appr. 5-10% increase) would result in a 9.1% drop in the number of births next year. CO pollution levels also have a significant although very small effect. If CO pollution concentration increases by 0.1 (appr. 1.5% increase) the number of births next year fall by 0.1%. The CO point estimate is significant at the 5% level.

Table 6: Desriptive statistics of the NUTS3 level variables.

Variable	Observations	Mean	Std. dev.	Min	Max
Birth rate	7,842	0.0535	0.0110	0.0217	0.1674
PM_{10}	7,731	22.8	11.9	1.6	137.6
SO_2	5,319	4.9	7.7	0.0	245.6
O3	7,098	52.0	10.8	5.3	111.6
NO_2	7,749	19.6	10.1	0.2	118.2
NOX	5,688	34.0	26.4	0.2	214.4
CO	3,866	4.1	46.7	0.0	1,225.9
C_6H_6	2,542	1.2	0.9	0.0	9.4
NO	5,057	11.0	11.7	0.0	254.0
Pb	2,225	0.7	15.4	0.0	539.8
$PM_{2.5}$	5,590	13.5	6.2	0.0	85.0
Temperature	9,200	10.9	2.8	0.5	20.9
Wind speed	9,200	3.0	0.9	1.0	5.7
HDD	7,456	2,620.2	867.2	266.6	6836.6
CDD	7,456	91.4	123.5	0.0	812.2
GDP (euro)	7,343	25,234	17,294	1,900	191,900
GDP (PPS)	7,343	25,404	13,403	4,400	173,800

Figure 1: Average birth rate in NUTS3 regions.

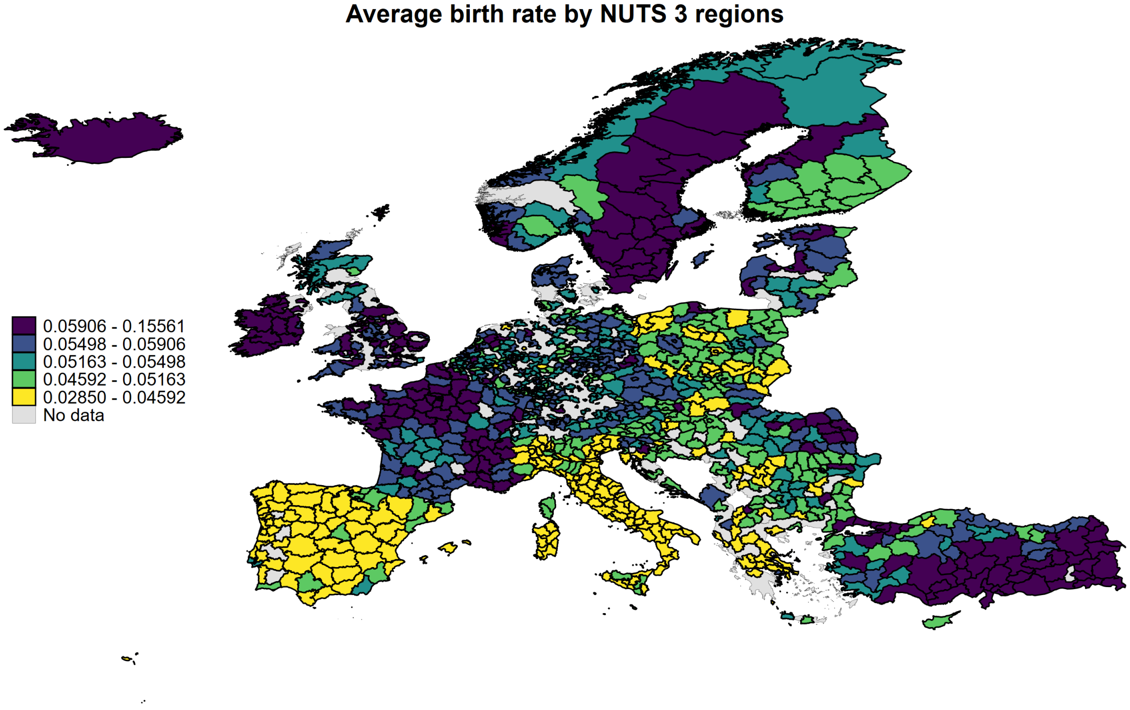

Figure 2: Average PM_{10} pollution in NUTS3 regions.

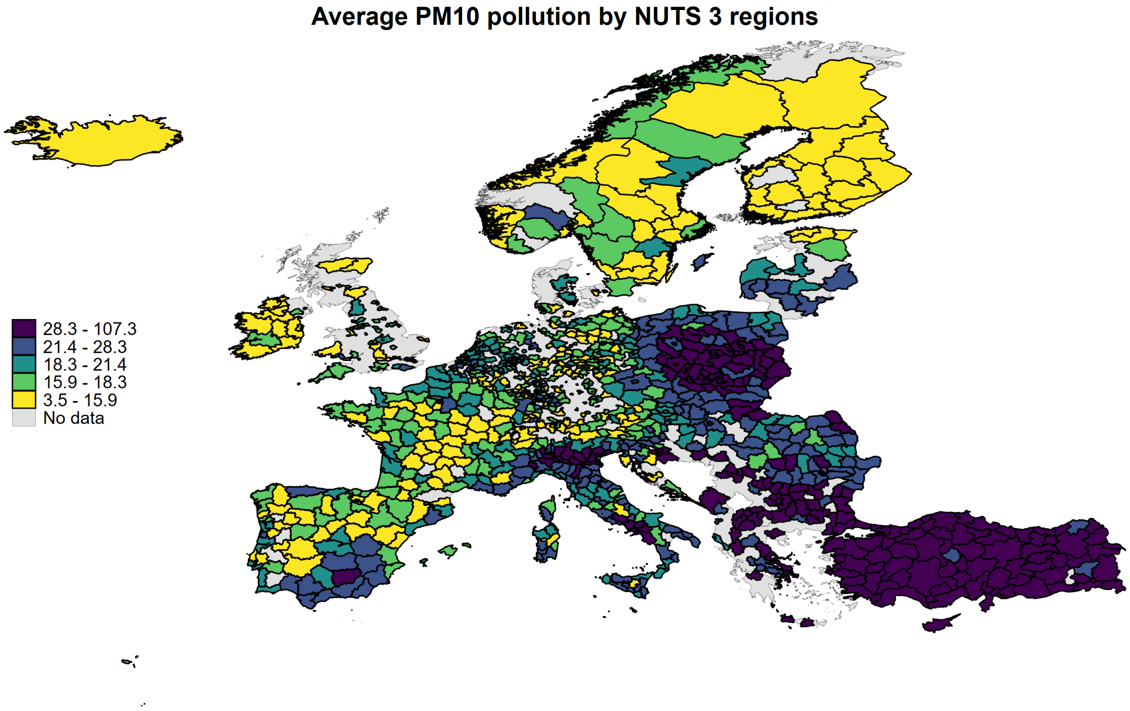

Table 7: Means of variables by year.

year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
PM_{10}	23.21	22.20	25.06	23.48	23.51	23.31	21.42	20.56
SO_2	4.47	4.60	5.79	5.10	4.93	4.96	4.71	4.79
O_3	52.13	48.84	52.53	49.53	52.11	54.59	54.02	52.08
NO_2	21.25	20.44	20.81	20.61	20.29	19.25	18.87	16.03
NOX	37.27	36.63	37.01	37.51	34.42	32.64	32.25	26.72
CO	0.80	13.87	14.21	2.84	0.47	0.42	0.62	0.41
$\mathrm{C}_6\mathrm{H}_6$	1.33	1.20	1.24	1.24	1.21	1.16	0.99	0.95
NO	11.62	12.35	13.02	13.31	12.08	10.21	9.64	7.61
Pb	0.23	2.24	1.37	1.53	0.20	0.02	0.07	0.02
$\mathrm{PM}_{2.5}$	15.18	14.33	14.37	13.63	13.68	13.75	12.26	11.85

After controlling for these effects, the rest of the pollutants do not have a significant effect on fertility in either of the specifications.

Table 8: The effect of mean ambient pollution concentrations on birth rates

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Factor (PM2.5, PM10) (t-1)	-0.027***	-0.027***	-0.026***	-0.091***
	[0.008]	[0.009]	[0.009]	[0.025]
SO2 (t-1)	-0.002	-0.003	-0.003	0.017
	[0.002]	[0.002]	[0.002]	[0.013]
Factor (NO2, NOX, O3) (t-1)	0.003	0.005	0.004	-0.060
	[0.007]	[0.007]	[0.007]	[0.039]
NO (t-1)	0.000	-0.000	0.000	0.002
	[0.000]	[0.000]	[0.000]	[0.003]
CO (t-1)	-0.00006***	-0.00006***	-0.00006***	-0.0001**
	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t-2)		-0.006	-0.007	-0.038
		[0.005]	[0.005]	[0.026]
SO2 (t-2)		0.001*	0.001*	0.010
		[0.000]	[0.000]	[0.013]
Factor (NO2, NOX, O3) (t-2)		0.004	0.004	-0.058
		[0.005]	[0.005]	[0.037]
NO (t-2)		-0.000	-0.000	0.000
		[0.000]	[0.000]	[0.002]
CO (t-2)		-0.000**	-0.000**	0.000
		[0.000]	[0.000]	[0.000]
GDP			0.000***	0.000***
			[0.000]	[0.000]
Constant	-2.954***	-2.953***	-3.075***	
	[0.006]	[0.007]	[0.038]	
Model	OLS	OLS	OLS	2SLS
Observations	2,800	2,800	2,800	2,800
F test	0.000	0.000	0.000	0.000

Robust standard errors in brackets. The standard errors are clustered at the NUTS 3 region level. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Instrumental variables estimates for various measures of ambient pollution

	(1) Mean	(2) Maximum	(3) Days D10	(4) Days D9-10	(5) Days D8-10
Factor (PM2.5, PM10) (t-1)	-0.091***	0.005	-0.072***	-0.092***	-0.081***
	[0.025]	[0.013]	[0.028]	[0.027]	[0.025]
SO2 (t-1)	0.017	0.000	0.002***	0.000*	0.000
	[0.013]	[0.001]	[0.001]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-1)	-0.060	0.005	-0.007	0.002	0.021
	[0.039]	[0.027]	[0.055]	[0.058]	[0.054]
NO (t-1)	0.002	-0.000	-0.001	0.000	0.000
	[0.003]	[0.000]	[0.001]	[0.000]	[0.000]
CO (t-1)	-0.000123**	-0.0000512*	-0.000133	-0.000227*	-0.000254**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t-2)	-0.038	0.000	-0.054**	-0.042	-0.046*
	[0.026]	[0.013]	[0.024]	[0.030]	[0.024]
SO2 (t-2)	0.010	-0.001	0.000	-0.000	-0.000
	[0.013]	[0.001]	[0.001]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-2)	-0.058	-0.057**	-0.049	0.004	0.009
	[0.037]	[0.022]	[0.040]	[0.033]	[0.034]
NO (t-2)	0.000	0.000	0.001**	0.000	0.000
	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
CO (t-2)	0.000	-0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
GDP	0.000***	0.000***	0.000***	0.000***	0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Model	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	2,800	2,800	2,927	2,927	2,927
F test	0.000	0.000	0.000	0.000	0.000

Robust standard errors in brackets. The standard errors are clustered at the NUTS 3 region level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

4.3. Robustness checks

Our first robustness check is to include concentration indices other than mean, to see whether our results are robust to these different measures of pollutant concentrations. In Table 9 we report the 2SLS estimation results for the effect of various measures of the ambient air pollution concentrations. In Col. 1 we repeat the results on the mean pollution values (Col. 4 of Table 8) for comparison. Col. 2 reports the point estimates for the maximum values of the concentrations. It is not surprising that these are insignificant because maximum values typically have large standard errors. In this specification, the effect of NO_2 factor becomes significant at the 5%. In Cols 3 to 5, we report the effect of days in the highest concentration decile (D10), days in deciles 9 and 10 (D9-10), and days in deciles 8 to 10 (D8-10). These results are very similar to the main regression results. Having 1 more day in a month on average when the $PM_{2.5}$ - PM_{10} pollution concentration is in the highest decile, would decrease fertility by 7.2%, and this result is significant at the 1% level.

In the second robustness check, we include placebo treatment variables, the future values of pollutant concentrations in the 2SLS regressions. These future values should not have a significant effect on the birth rates, and the estimated effects of the past values should not differ from the baseline specification. We report the results of the placebo treatment in Table 10, and we see exactly what we expected.

Next, we extend the measures of pollution concentration already reported in Table 9. First, we run a regression where we include the number of average days per month when the pollution concentration reached the highest decile (D1), we go on with the highest two deciles (D2). We expect that the more deciles we add to the measurement, the lower effect this index should have on the birth rate. Lastly, extend the deciles to ten which ultimately includes all days and should not have a significant effect on birth rates. The results are reported in Table 11 and the results align with our expectations.

Table 10: Placebo test for future pollution concantrations

	(1) Mean	(2) Maximum	(3) Days D10	(4) Days D9-10	(5) Days D8-10
Factor (PM2.5, PM10) (t-1)	-0.075***	0.010	-0.081***	-0.094***	-0.088***
	[0.026]	[0.016]	[0.030]	[0.034]	[0.026]
SO2 (t-1)	-0.001	-0.000	0.001	0.000	0.000
	[0.012]	[0.001]	[0.001]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-1)	-0.005	-0.009	-0.014	0.082	0.102*
	[0.035]	[0.025]	[0.043]	[0.074]	[0.061]
NO (t-1)	0.003	-0.000	-0.000	0.000	0.000
	[0.002]	[0.000]	[0.001]	[0.001]	[0.000]
CO (t-1)	-0.000*	-0.000	0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t-2)	-0.005	0.011	-0.023	-0.004	-0.031
	[0.027]	[0.013]	[0.023]	[0.031]	[0.024]
SO2 (t-2)	0.001	-0.000	0.000	-0.000	-0.000
	[0.005]	[0.001]	[0.001]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-2)	-0.032	-0.056***	-0.077**	-0.011	-0.005
	[0.030]	[0.018]	[0.037]	[0.042]	[0.043]
NO (t-2)	0.003	0.000	0.001***	0.000	0.000
	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]
CO (t-2)	-0.000	-0.000	0.000	0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t+1)	0.035	-0.002	-0.051	-0.044	-0.022
	[0.033]	[0.012]	[0.049]	[0.048]	[0.029]
SO2 (t+1)	-0.005	-0.000	0.000	0.000	0.000
	[0.005]	[0.001]	[0.000]	[0.001]	[0.000]
Factor (NO2, NOX, O3) (t+1)	-0.044	0.010	-0.045	0.013	0.004
	[0.042]	[0.031]	[0.081]	[0.063]	[0.048]
NO (t+1)	0.002	0.000	0.001	-0.000	-0.000
	[0.003]	[0.000]	[0.001]	[0.000]	[0.000]
CO (t+1)	-0.000	0.000	0.000	0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
GDP	0.000***	0.000***	0.000***	0.000***	0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Model	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	2,462	2,462	2,584	2,584	2,584
F-statistic	0.000	0.000	0.000	0.000	0.000

Robust standard errors in brackets. The standard errors are clustered at the NUTS 3 region level. *** p<0.01, ** p<0.05, * p<0.1

Table 11: The effect of days spent in the highest deciles of pollution concentration on birth rates

	(1) D1	(2) D2	(3) D3	(4) D4	(5) D5	(6) D6	(7) D7	(8) D8	(9) D9	(10) D10
Factor (PM2.5, PM10) (t-1)	-0.072***	-0.092***	-0.081***	-0.065***	-0.056***	-0.061***	-0.046**	-0.049**	-0.035	-0.060*
	[0.028]	[0.027]	[0.025]	[0.025]	[0.021]	[0.023]	[0.023]	[0.023]	[0.025]	[0.034]
SO2 (t-1)	0.002***	0.000*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000**
	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-1)	-0.007	0.002	0.021	-0.012	-0.012	-0.011	-0.025	-0.018	-0.026	-0.017
	[0.055]	[0.058]	[0.054]	[0.061]	[0.053]	[0.049]	[0.042]	[0.040]	[0.038]	[0.031]
NO (t-1)	-0.001	0.000	0.000	0.000	0.000	0.000*	0.000**	0.000	0.000	0.000
	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
CO (t-1)	-0.000	-0.000*	-0.000**	-0.000**	-0.000**	-0.000*	-0.000	-0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t-2)	-0.054**	-0.042	-0.046*	-0.037	-0.026	-0.038	-0.009	-0.009	0.017	0.012
	[0.024]	[0.030]	[0.024]	[0.026]	[0.027]	[0.026]	[0.024]	[0.027]	[0.023]	[0.018]
SO2 (t-2)	0.000	-0.000	-0.000	-0.000	-0.000	0.000*	0.000	0.000	-0.000	0.000
	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Factor (NO2, NOX, O3) (t-2)	-0.049	0.004	0.009	0.013	0.018	0.038	0.011	0.010	-0.003	-0.005
	[0.040]	[0.033]	[0.034]	[0.039]	[0.037]	[0.043]	[0.031]	[0.034]	[0.031]	[0.017]
NO (t-2)	0.001**	0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
CO (t-2)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
GDP	0.000***	0.000**	0.000**	0.000**	0.000**	0.000**	0.000***	0.000***	0.000***	0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Model	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	2,927	2,927	2,927	2,927	2,927	2,927	2,927	2,927	2,927	2,927
F test	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in brackets. The standard errors are clustered at the NUTS 3 region level. *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the logarithm of birth rate in year t in NUTS 3 region r.

Table 12: Heterogeneity results using 2SLS regressions

	(1) High pollution	(2) Low pollution	(3) High GDP	(4) Low GDP
Factor (PM2.5, PM10) (t-1)	-0.121***	-0.068**	-0.058*	-0.099***
	[0.045]	[0.029]	[0.034]	[0.034]
SO2 (t-1)	0.017	-0.007	-0.000	0.000
	[0.011]	[0.014]	[0.015]	[0.010]
Factor (NO2, NOX, O3) (t-1)	-0.006	-0.025	-0.001	-0.013
	[0.060]	[0.032]	[0.027]	[0.084]
NO (t-1)	0.001	0.004	-0.002	0.003
	[0.002]	[0.003]	[0.002]	[0.006]
CO (t-1)	-0.000	-0.000**	-0.000	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]
Factor (PM2.5, PM10) (t-2)	-0.052	-0.017	-0.059**	-0.022
	[0.039]	[0.035]	[0.028]	[0.026]
SO2 (t-2)	0.007	0.012	-0.017	0.002
	[0.012]	[0.012]	[0.013]	[0.006]
Factor (NO2, NOX, O3) (t-2)	0.026	-0.064**	-0.015	0.046
	[0.032]	[0.032]	[0.035]	[0.043]
NO (t-2)	0.001	0.000	0.003*	-0.005
	[0.001]	[0.003]	[0.002]	[0.004]
CO (t-2)	0.000	0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
GDP	0.000	0.000***	0.000***	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
Model	2SLS	2SLS	2SLS	2SLS
Observations	964	1,836	1,896	904
F test	0.051	0.000	0.000	0.001

Robust standard errors in brackets. The standard errors are clustered at the NUTS 3 region level. *** p<0.01, ** p<0.05, * p<0.1

High pollution: average PM Factor in the observation period is higher than the median of all regions. Low pollution: average PM Factor in the observation period is lower than the median of all regions. High GDP: average GDP in the observation period is higher than the median of all regions. Low GDP: average GDP in the observation period is lower than the median of all regions.

4.4. Heretogeneity analysis

Finally, we present a heterogeneity analysis, and the results are reported in Table 12. We find intuitive results. A similar increase in the pollution concentration deteriorates birth rates if the pollution levels were already high. On the other hand, in countries with lower GDPs pollution decreases birth rates to a much greater extent. This result is probably due to the higher quality of health services or to the generally better health status of the population.

5. Discussion

In this paper, we examined how different types of ambient pollutants affect birth rates in the European Union. We have found that it is the particulate matter concentrations, specifically $PM_{2.5}$ and PM_{10} that have a significant effect on birth rates. After controlling for these effects, the rest of the pollutants were found to exert an insignificant effect on fertility. The PM coefficient is significant at the 1% level, and it suggests that an increase in the levels of $PM_{2.5}$ - PM_{10} pollution concentration by 1 (appr. 5-10% increase) would result in a 9.1% drop in the number of births next year.

Our heterogeneity analysis shows that air pollution concentrations have a much larger effect in countries with lower GDP. In countries with lower than median GDP, the point estimate is -0.099 and significant at 1%, whereas, in countries with high GDP, the point estimate is -0.058 and insignificant. These results point to the importance of moderating factors such as the quality and accessibility of health services or the general health status of the population.

References

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal*, 126(591): 257–280. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12273.

Benmarhnia, Tarik, Prashant Bharadwaj, and Mauricio Romero. "Using Instrumental Variables under Partial Observability of Endogenous Variables for Assessing Effects of Air Pollution on Health."

Bové, Hannelore, Eva Bongaerts, Eli Slenders, Esmée M. Bijnens, Nelly D. Saenen, Wilfried Gyselaers, Peter Van Eyken, Michelle Plusquin, Maarten B. J. Roeffaers, Marcel Ameloot, and Tim S. Nawrot. 2019. "Ambient black carbon particles reach the fetal side of human placenta." *Nature Communications*, 10(1): 3866. Number: 1 Publisher: Nature Publishing Group.

Conforti, Alessandro, Marika Mascia, Giuseppina Cioffi, Cristina De Angelis, Giuseppe Coppola, Pasquale De Rosa, Rosario Pivonello, Carlo Alviggi, and Giuseppe De Placido. 2018. “Air pollution and female fertility: a systematic review of literature.” *Reproductive biology and endocrinology: RB&E*, 16(1): 117.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. “The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction.” *American Economic Review*, 109(12): 4178–4219.

ECA. 2018. “Air pollution: Our health still insufficiently protected.” European Court of Auditors Special Report 23/2018.

Faiz, Ambarina S., George G. Rhoads, Kitaw Demissie, Lakota Kruse, Yong Lin, and David Q. Rich. 2012. “Ambient air pollution and the risk of stillbirth.” *American Journal of Epidemiology*, 176(4): 308–316.

Frutos, Víctor, Mireia González-Comadrán, Ivan Solà, Benedicte Jacquemin, Ramón Carreras, and Miguel A. Checa Vizcaíno. 2015. “Impact of air pollution on fertility: a systematic review.” *Gynecological Endocrinology*, 31(1): 7–13.

Hajdu, Tamás, and Gábor Hajdu. 2022. “Temperature, climate change, and human conception rates: evidence from Hungary.” *Journal of Population Economics*, 35(4): 1751–1776.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. “Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health.” *Massachusetts Institute of Technology Press*. Accepted: 2018-03-02T19:11:39Z Publisher: MIT Press.

Legro, Richard S., Mark V. Sauer, Gilbert L. Mottla, Kevin S. Richter, Xian Li, William C. Dodson, and Duanping Liao. 2010. “Effect of air quality on assisted human reproduction.” *Human Reproduction (Oxford, England)*, 25(5): 1317–1324.

Levine, Hagai, Niels Jørgensen, Anderson Martino-Andrade, Jaime Mendiola, Dan Weksler-Derri, Irina Mindlis, Rachel Pinotti, and Shanna H. Swan. 2017. “Temporal trends in sperm count: a systematic review and meta-regression analysis.” *Human Reproduction Update*, 23(6): 646–659.

Mohallem, Soraya Vecci, Débora Jã de Araújo Lobo, Célia Regina Pesquero, João Vicente Assunção, Paulo Afonso de Andre, Paulo Hilário Nascimento Saldiva, and Marisa Dolhnikoff. 2005. “Decreased fertility in mice exposed to environmental air pollution in the city of Sao Paulo.” *Environmental Research*, 98(2): 196–202.

Mohorovic, Lucijan, Oleg Petrovic, Herman Haller, and Vladimir Micovic. 2010. “Pregnancy loss and maternal methemoglobin levels: an indirect explanation of the association of environmental toxics and their adverse effects on the mother and the fetus.” *International Journal of Environmental Research and Public Health*, 7(12): 4203–4212.

Qian, Hong, Qiaoqiao Xu, Wenkai Yan, Yun Fan, Zhi Li, Chengzhe Tao, Feng Zhang, and Chuncheng Lu. 2022. "Association between exposure to ambient air pollution and semen quality in adults: a meta-analysis." *Environmental Science and Pollution Research*, 29(7): 10792–10801.

Schwartz, Joel, Elena Austin, Marie-Abele Bind, Antonella Zanobetti, and Petros Koutrakis. 2015. "Estimating Causal Associations of Fine Particles With Daily Deaths in Boston." *American Journal of Epidemiology*, 182(7): 644–650.

Schwartz, Joel, Marie-Abele Bind, and Petros Koutrakis. 2017. "Estimating Causal Effects of Local Air Pollution on Daily Deaths: Effect of Low Levels." *Environmental Health Perspectives*, 125(1): 23–29.

Slama, Rémy, Sébastien Bottagisi, Ivo Solansky, Johanna Lepeule, Lise Giorgis-Allemand, and Radim Sram. 2013. "Short-term impact of atmospheric pollution on fecundability." *Epidemiology (Cambridge, Mass.)*, 24(6): 871–879.

WHO. 2019. "Ten threats to global health in 2019."

A1. Supplementary tables and figures

Table A.13: Number of data points before and after aggregating.

	raw data		**data after aggregating**	
	by station	**by station**	**by nuts3**	**by nuts3**
Pollutant	**hour**	**day**	**day**	**year**
C_6H_6	31 256 798	224 539	754 028	2 542
CO	72 715 222	10 248	1 331 598	3 866
NO	142 830 222	7 672	1 771 263	5 058
NO_2	231 221 335	52 483	2 742 982	7 749
NOX as NO_2	136 133 784	10 460	1 988 002	5 688
O_3	163 347 466	2 190	2 510 812	7 098
Pb in PM_{10}	26 739	1 047 140	474 413	2 225
PM_{10}	150 259 482	3 542 713	2 718 911	7 731
$PM_{2.5}$	71 888 620	1 656 595	1 893 156	5 590
SO2	122 923 200	66 823	1 847 375	5 319

Table A.14: EU air quality standards

Pollutant	**Concentration limit (CL)**	**Averaging period**
Fine particles ($PM_{2.5}$)	20 µg/m^3	1 year
Sulphur dioxide (SO_2)	125 µg/m^3	24 hours
Nitrogen dioxide (NO_2)	40 µg/m^3	1 year
Particulate matter (PM_{10})	50 µg/m^3	24 hours
Lead (Pb)	0.5 µg/m^3	1 Year
Carbon monoxide (CO)	10 mg/m^3	Maximum daily 8 hour mean
Benzene (C_6H_6)	5 µg/m^3	1 year
Ozone (O_3)	120 µg/m^3	Maximum daily 8 hour mean