Human Sex Ratio at Birth: Mythical and Observable Trends

Yanan Long[1,2], **Qi Chen**[3], **Henrik Larsson**[3,4], **Andrey Rzhetsky**[2,5*]

***For correspondence:**
andrey.rzhetsky@uchicago.edu (AR)

[1]Department of Chemistry, The University of Chicago, Chicago, IL, 60637, United States of America; [2]Institute of Genomics and Systems Biology, The University of Chicago, Chicago, IL, 60637, United States of America; [3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE-171 76, Sweden; [4]School of Medical Sciences, Örebro University, Örebro, SE-701 82, Sweden; [5]Departments of Medicine and Human Genetics, The University of Chicago, Chicago, IL, 60637, United States of America

Abstract The human sex ratio at birth (SRB) is defined as the ratio between the number of newborn boys to the total number of newborns per time unit. It is, typically, slightly greater than $\frac{1}{2}$ (more boys than girls) and fluctuates over time. In this study, we sought to "myth-check" previously reported associations (and test new hypotheses) using variants of mixed-effect regression analyses and time-series models on two very large electronic health record datasets, representing the populations in the United States and Sweden, respectively. Our results revealed that neither dataset supported models in which the SRB changed seasonally or in response to variations in ambient temperature, and that an increased level of a diverse array of pollutants were associated with lower SRBs. Moreover, we found that increased levels of industrial and agricultural activity, which served as proxies for water pollution, were also associated with lower SRBs.

Introduction

Because human male gametes bearing X or Y chromosomes are equally frequent (being produced by meiosis symmetrically partitioning two sex chromosomes), and because female gametes bear only X chromosomes, one would expect that the SRB value at conception is exactly $\frac{1}{2}$. A recent study (***Orzack et al., 2015***) using the fluorescent *in situ* hybridization technique to determine the sex of embryos conceived through *in vitro* fertilization estimated that sex ratio at conception is indeed statistically indistinguishable from one half, 0.502 (***Orzack et al., 2015***). Nonetheless, historically observed SRB values have been, typically, significantly different from $\frac{1}{2}$ (***Bruckner and Catalano, 2018***), which means that SRBs must fluctuate due to the differential survival of each sex's embryos and fetuses (***Bruckner and Catalano, 2018***). One possible mechanism for higher SRBs is that more female pregnancies terminate before they are clinically detectable. These excess female losses occur primarily during the first and early-second trimesters, whereas excess male losses occur during the late-second and third trimesters (***Bruckner and Catalano, 2018***). In the late-second and third trimesters, SRB again reduces due to that male-biased mortality.

Previous literature suggests that the SRB can fluctuate with time (***Bruckner and Catalano, 2018***) and is putatively associated with a number of environmental factors, such as chemical pollution, catastrophic events that exert stress on pregnant women (*e.g.*, terrorist attacks and earthquakes), radiation, weather parameters, and even seasons of conception (Table 1). While there are multiple studies in which the positive associations between air pollution and spontaneous abortion have been observed (***Leiser et al., 2019***; ***Zhang et al., 2019***), most of those conclusions were drawn based on analyses of relatively small samples (Table 1). In this study, we harnessed the power of two very

large datasets: MarketScan's insurance claim data (***Hansen, 2017***) in the United States (more than 100 million unique patients and more than three million unique newborns recorded between 2003 to 2011), and Sweden's birth registry data (~ 3.2 million newborns from 1983 to 2013) (***Emilsson et al., 2015***). This is the first systematic investigation of numerous chemical pollutants and other environmental factors using large datasets from two continents.

Methods

Data

The IBM Health MarketScan data (***Hansen, 2017***) represents $104,565,671$ unique individuals and $3,134,062$ unique live births. The Swedish Health Registries (***Emilsson et al., 2015***) record health statistics for over ten million individuals, and $3,260,304$ unique live births. We juxtaposed time-stamped birth events in the two countries with exogenous factor measurements retrieved from the US National Oceanic and Atmospheric Administration, the US Environmental Protection Agency, the Swedish Meteorological and Hydrological Institute, and Statistics Sweden. We used a subset of the MarketScan data (***Hansen, 2017***) which contained information about live births during the period of 2003 to 2011 with county information encoded in FIPS (Federal Information Processing Standards) codes and a family link profile indicating the composition of the families contained in the dataset. The date, geographic distribution, and the mothers of the newborns can be directly extracted from these datasets. For environmental factors, we used the Environmental Quality Index (EQI) data compiled by the United States Environmental Protection Agency (EPA) (***Lobdell et al., 2011***; ***Messer et al., 2014***).

Cluster Analysis

In order to simplify subsequent analyses, we first reduced the dimensionality of the EQI dataset by performing hierarchical clustering analysis on the matrix of Spearman's rank correlation coefficients (ρ) with the Ward's method. We then used the R-language (***R Core Team, 2017***) package `pvclust` (***Suzuki and Shimodaira, 2006***) to minimize the total within-cluster variance (***Legendre and Legendre, 2012***). The resulting dendrogram and list of factors can be found in the SI. Each cluster contains at least two factors and is represented by the mean of all the elements in the cluster.

Regression Analysis

We used Bayesian logistic regression with random effects, implemented in the R-language package `brms` (***Bürkner, 2017***) and the No-U-Turn sampler (NUTS) (***Hoffman and Gelman, 2014***). The model for the jth factor (predictor) is given as follows:

$$\text{logit}\left(p_j\right) = \log\left(\frac{p_j}{1-p_j}\right) = \alpha_{[k]j} + \boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}_j, \tag{1}$$

where p_j is the probability that a newborn is male, $\mathbf{x}_j$ is the vector representing the jth factor, $\boldsymbol{\beta}_j$ their coefficients, and $\alpha_{[k]j}$ the intercept for the kth group-level, representing states or counties in the US, and *kommuner* (municipalities) or *län* (counties) in Sweden, if applicable. The group-level effect is modeled for a single random effect by (***Gelman and Hill, 2006***)

$$\alpha_j = \mu_a + \eta_j, \tag{2}$$

$$\eta_j \sim \mathcal{N}\left(0, \sigma_\eta^2\right) \tag{3}$$

and for two random effects, representing e.g. state- and county-specific effects, by

$$\alpha_j = \mu_a + \eta_j + \nu_j, \tag{4}$$

$$\eta_j \sim \mathcal{N}\left(0, \sigma_\eta^2\right), \tag{5}$$

$$\nu_j \sim \mathcal{N}\left(0, \sigma_\nu^2\right), \tag{6}$$

where η_j and ν_j are independent of each other and for all j. Moreover, we partitioned the independent variables into septiles, so that $\boldsymbol{\beta}_j \in \mathbb{R}^6$, with one regression coefficient for each of the six septiles other than the first, which was chosen as the reference level (***Khan et al., 2019***).

We applied logistic regression in two ways. First, to assess the effect of environmental factors, we regressed the septiles of each of the individual factors against the SRB, with each sample point representing a county. Therefore, each septile, aside from the baseline, has a coefficient. Second, to test whether a mother's diagnostic history (DX) affected the SRB, we regressed a DX's indicator variables against the SRB, with each sample point representing a newborn/mother pair. For model selection in both cases, we performed repeated (10 times) 10-fold cross-validation and calculated the information criterion relative to the null model (where $\mathbf{x}_j = \mathbf{0}$, i.e., the model comprised solely of the intercept). We computed the average difference in information criterion ($\Delta\mathrm{IC}$) and standard error (SE) for each factor, and used the Benjamini–Yekutieli method to account for multiple comparisons (***Benjamini et al., 2001***).

Univariate Time-series Analysis

To assess the effect of one-off, stressful events on the SRB, we used two different time series techniques. We first fitted seasonal univariate autoregressive, integrated, moving average (sARIMA) models using the Box-Jenkins method (***Box et al., 2015***) in conjunction with monthly (28-day periods) and weekly live birth data up to the event and then performed an out-of-sample prediction. A sARIMA model is given as follows:

$$\underbrace{\left(1 - \sum_{j=1}^{P} \Phi_j L^j\right)}_{\text{sAR}} \underbrace{\left[\left(1 - L^S\right)^D\right]}_{\text{sI}} \underbrace{\left(1 - \sum_{j=1}^{p} \phi_j L^j\right)}_{\text{AR}} \underbrace{\left[(1 - L)^d\right]}_{\text{I}} y_t = \underbrace{\left(1 + \sum_{j'=1}^{Q} \Theta_{j'} L^{j'}\right)}_{\text{sMA}} \underbrace{\left(1 + \sum_{j'=1}^{q} \theta_{j'} L^{j'}\right)}_{\text{MA}} \varepsilon_t, \quad (7)$$

where AR indicates the auto-regression term, I the integration term, and MA the moving average term (an s before any of the above stands for seasonal), y_t the observed univariate time series of interest, L the lag operator such that $L(y_t) = y_{t-1}$, ε's white noises, $S \geqslant 2$ the degree of seasonality (i.e., the number of seasonal terms per year), and ϕ's, θ's, Φ's and Θ's are model parameters to be estimated. We used the R package's `forecast` (***Hyndman and Khandakar, 2008***; ***Hyndman et al., 2019***) `auto.arima` function to fit the data, which performs a step-wise search on the (p, d, q, P, D, Q) hyperparameter space and compares different models by using the Bayesian Information Criterion (BIC) (***Schwarz, 1978***). We confirmed the optimal models' goodness-of-fit using the Breusch-Godfrey test on the residuals, which tests for the presence of autocorrelation up to degree S (***Breusch, 1978***; ***Godfrey, 1978***; ***Hayashi, 2000***). On the other hand, we fitted the same data as above to Bayesian structural time series (BSTS) models, which are state space models given in the general form by (***Scott and Varian, 2013***):

$$y_t = Z_t^{\intercal} \alpha_t + \varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, H_t) \quad (8)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim \mathcal{N}(0, Q_t), \quad (9)$$

where y_t is the observed time series and α_t the *unobserved* latent state. In particular, we used the local linear trend model with additional seasonal terms (***Murphy, 2012***; ***Scott and Varian, 2013***):

$$y_t = \mu_t + \tau_t + \varepsilon_t \quad (10)$$

$$\mu_t = \mu_t + \delta_{t-1} + u_t \quad (11)$$

$$\delta_t = \delta_{t-1} + v_t \quad (12)$$

$$\tau_t = -\sum_{s=1}^{S-1} \tau_{t-s} + w_t. \quad (13)$$

Under this formulation, we have $\eta_t = \begin{bmatrix} u_t & v_t & w_t \end{bmatrix}$ and Q_t is a t-invariant block diagonal matrix with diagonal elements σ_u^2, σ_v^2 and σ_w^2. Furthermore, we write $\alpha_t = \begin{bmatrix} \mu_t & \delta_t & \tau_{t-S+2} & \cdots & \tau_t \end{bmatrix}$, which implies

that both Z_t and T_t are t-invariant matrices of 0's and 1's such that Equations 10–13 hold. We used the R package `CausalImpact` (***Brodersen et al., 2015***), which uses the R package `bsts` (***Scott, 2019***) as backend, to fit the data.

Correlation and Causality

To test whether the SRB was effected by ambient temperature, we grouped daily SRB data and temperatures into 91-day (13-week) periods and calculated the Pearson correlation coefficient (r) between each SRB and ambient temperature. We then performed the Student's t-test for the null hypothesis that the true correlation is 0. Furthermore, we fitted the SRB/temperature pair to a vector autoregression (VAR) model for a maximum lag order of 4 (52 weeks), using the BIC as the metric for model selection, and then tested for the null hypothesis of the non-existence of Granger causality using the F-test (***Granger, 1969***).

Results

We start by describing the negative results concordant across the two datasets. Our model selection results rejected the whole spectrum of models that postulate periodic, annual SRB changes (***Lerchl, 1998***; ***Melnikov and Grech, 2003***). For both US and Swedish datasets, the best-fitting sARIMA model described the SRB as lacking seasonality throughout the year. Similarly, when we tested the claim that ambient temperatures during conception affect the SRB (***Catalano et al., 2008***; ***Helle et al., 2009***; ***Fukuda et al., 2014***; ***Dixson et al., 2011***), we found that neither dataset supported this association, as both the Student's t-test and the F-test concluded that the SRB was independent of ambient temperature measurements (Table S7).

A comparison of each dataset's environmental measurements revealed that Sweden enjoyed both lower variations and lower mean values of measured concentrations of substances in the air. Unfortunately, the Swedish dataset also provided fewer measured pollutants, which made our cross-country analysis more difficult. Figure 1**A** shows a comparison of pollutant concentration distributions in both countries. The US environmental measurements dataset presented its own difficulty: Many pollutants appeared highly collinear in their spatial variation. To circumvent this, we performed a cluster analysis on the environmental factors, subdividing them into 26 clusters (see the *Cluster Analysis* section in *Methods* and Table 4). All pollutants within the same cluster were highly correlated, while the correlation between distinct clusters was smaller, allowing for useful inferences of associations between SRB changes and environmental states.

Using the US dataset, we were able to validate a number of previous studies' findings regarding the association between the SRB and exogenous factors (Table 2). Specifically, our data suggests that PCBs (polychlorinated biphenyls), aluminium in air, chromium in water, and total mercury appear to drive the SRB up, while lead in soil appears to be associated with a decreased SRB. Meanwhile, we refute a number of previous reports (these are indicated with a dash in the second column in Table 2). Finally, we also found a number of new environmental associations that have not been reported before (see Table 3, Figure 1 **B–M**, and Figure 2). Figure 1 **B–M** shows that increased pollutant levels appear to be associated with both increased and decreased SRB values (Plates **E**,**F**,**H**,**I**, and **J**, and the remaining Plates, respectively).

The geographic distribution of these pollutants varies remarkably, as seen in Figure 2. For example, lead in land (Figure 2**B**) appears to be enriched in the northeast, southwest, and mid-east US, but not in the south. Hydrazine (Figure 2**F**) appears to follow capricious, blotch-like shapes in the eastern US, each blotch likely centered at a factory emitting this pollutant. Total mercury deposition in water (Figure 2**E**) mostly affects eastern US states with the heaviest load in the northeastern states. It is this variability in the distribution of various substances in the environment that makes teasing out individual associations possible.

Finally, when we tested links between two stressful events in the US (Hurricane Katrina and the Virginia Tech shooting) and the SRB using seasonal autoregressive integrated moving-average (sARIMA) models and state-space models (SSMs) (see the *Univariate time-series analysis* section

in *Methods*), we were able to identify significant associations only in the case of the Virginia Tech shooting – the SRB was lower than expected 34 weeks after the event (see Figures S5(c) and S6(c) as well as Tables S5(c) and S6(c)).

Discussion

The literature is replete with theories striving to explain SRB changes in terms of evolutionary stresses and benefits. For example, one such theory, named the Trivers-Willard hypothesis after the researchers who proposed it (***Trivers and Willard, 1973***), states that the SRB fluctuates due to environmental conditions changing from favourable to unfavourable and back. Under favourable conditions, a male descendent is postulated to generate many more offspring than a female, and vice versa under unfavourable circumstances. Under this theory, parents maximize the number of grandchildren, thereby regulating the SRB to this end. The empirical support for this hypothesis is controversial (***Keller et al., 2001***; ***Lynch et al., 2018***; ***Song, 2018***), and the theory is vague when it comes to elucidating the underlying biological mechanisms. Taken as whole, our results appear to cast serious doubt on the validity of the Trivers-Willard hypothesis.

Because the average observed SRB is higher than $\frac{1}{2}$, we propose to explain these observations through the differential survival of either gametes or fetuses. Specifically, we hypothesize that female fetuses are subject to a higher risk of termination than male ones, which seems contrary to the theory that the SRB skews towards males (***Orzack et al., 2015***). To explain this inconsistency, some researchers have suggested that more female than male pregnancies terminate before they are clinically detectable, that is, excess female losses occur primarily during the first and early-second trimesters, whereas excess male losses occur during the late-second and third trimesters (***DiPietro and Voegtline, 2017***; ***Bruckner and Catalano, 2018***). In other words, the SRB starts at very close to $\frac{1}{2}$, then increases due to female-biased mortality, and finally reduces due to male-biased mortality. Because the physiology of male and female fetuses differs (***DiPietro and Voegtline, 2017***), it is conceivable that specific environmental stresses affect fetuses of different sexes in a sex-specific way.

Recent studies have uncovered a positive association between pollution and spontaneous abortion (***Leiser et al., 2019***; ***Zhang et al., 2019***), so our analyses presented here allow for more specific hypotheses regarding the potential chemicals that might cause preferential abortions of male or female fetuses. Fortunately, most of the putative positive associations identified in our study are experimentally testable, at least with animal models. However, given the large sample sizes of our analysis, one would expect that it is rather unlikely that the associations rejected in our study, such as those between SRB and ambient temperature and seasonality, as well as some environmental factors (see Tables 2 and 3), will be found as valid in other large datasets or laboratory studies.

Because the observed, statistically-significant SRB fluctuations have to be mediated by undetected spontaneous abortions, it is important to better understand the underlying mechanisms. The current literature is regretfully meagre regarding causal mechanisms behind the temporal fluctuations of the SRB as well as the plausible ramifications of the presence of environmental toxins and their possible remedies.

Finally, our study also highlights the need for more detailed and systematic measurements of environmental quality data – spatially, temporally, and across the chemical spectrum. For the US, the current measurements are available at a very coarse, county-level scale and are not repeated frequently enough. Ideally, the measurements would be taken frequently enough (say, daily or weekly) to enable the application of time-series methods (e.g. Granger causality or cointegration tests), which would help determine the causal structures of the time variation of exogenous factors on the SRB (***Samartsidis et al., 2018***). On the other hand, only a small subset of US environmental quality measurements were available in Sweden. As a result, our Swedish analyses lacks both the statistical power and breadth of measurements required to reproduce the kinds of positive,

statistical associations between SRB and individual environmental factors that we discovered in the US data.

Acknowledgments

We are grateful to E. Gannon and M. Rzhetsky for comments on earlier versions of this manuscript. This work was funded by the DARPA Big Mechanism program under ARO contract W911NF1410333, by National Institutes of Health grants R01HL122712, 1P50MH094267, and U01HL108634-01, and by a gift from Liz and Kent Dauten.

The authors declare that they have no competing financial interests.

References

Benjamini Y, Yekutieli D, et al. The control of the false discovery rate in multiple testing under dependency. The annals of statistics. 2001; 29(4):1165–1188.

Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.

Breusch TS. Testing for autocorrelation in dynamic linear models. Australian Economic Papers. 1978; 17(31):334–355.

Brodersen KH, Gallusser F, Koehler J, Remy N, Scott SL, et al. Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics. 2015; 9(1):247–274.

Bruckner TA, Catalano R. Selection in utero and population health: theory and typology of research. SSM-population health. 2018; 5:101–113.

Bürkner PC. brms: An R package for Bayesian multilevel models using STAN. Journal of Statistical Software. 2017; 80(1):1–28.

Catalano R, Bruckner T, Smith KR. Ambient temperature predicts sex ratios and male longevity. Proceedings of the National Academy of Sciences. 2008; 105(6):2244–2247.

Catalano RA. Sex ratios in the two Germanies: a test of the economic stress hypothesis. Human Reproduction. 2003; 18(9):1972–1975.

DiPietro JA, Voegtline KM. The gestational foundation of sex differences in development and vulnerability. Neuroscience. 2017; 342:4–20.

Dixson BJ, Haywood J, Lester PJ, Ormsby DK. Whatever the weather: ambient temperature does not influence the proportion of males born in New Zealand. PloS one. 2011; 6(9):e25064.

Emilsson L, Lindahl B, Köster M, Lambe M, Ludvigsson JF. Review of 103 Swedish Healthcare Quality Registries. Journal of internal medicine. 2015; 277(1):94–136.

Fukuda M, Fukuda K, Shimizu T, Nobunaga M, Mamsen LS, Andersen CY. Climate change is associated with male: female ratios of fetal deaths and newborn infants in Japan. Fertility and sterility. 2014; 102(5):1364–1370.

Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. Cambridge university press; 2006.

Godfrey LG. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. Econometrica: Journal of the Econometric Society. 1978; p. 1293–1301.

Granger CW. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society. 1969; p. 424–438.

Hansen L. The Truven Health MarketScan databases for life sciences researchers. Truven Health Ananlytics IBM Watson Health. 2017; .

Hayashi F. Econometrics. Princeton University Press. Section; 2000.

Helle S, Helama S, Lertola K. Evolutionary ecology of human birth sex ratio under the compound influence of climate change, famine, economic crises and wars. Journal of Animal Ecology. 2009; 78(6):1226–1233.

Hoffman MD, Gelman A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research. 2014; 15(1):1593–1623.

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F. forecast: Forecasting functions for time series and linear models; 2019, http://pkg.robjhyndman.com/forecast, r package version 8.5.

Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software. 2008; 26(3):1–22. http://www.jstatsoft.org/article/view/v027i03.

James WH, Grech V. A review of the established and suspected causes of variations in human sex ratio at birth. Early human development. 2017; 109:50–56.

Keller MC, Nesse RM, Hofferth S. The Trivers–Willard hypothesis of parental investment: no effect in the contemporary United States. Evolution and Human Behavior. 2001; 22(5):343–360.

Khan A, Plana-Ripoll O, Antonsen S, Brandt J, Geels C, Landecker H, Sullivan PF, Pedersen CB, Rzhetsky A. Environmental pollution is associated with increased risk of psychiatric disorders in the US and Denmark. PLoS biology. 2019; 17(8):e3000353.

Legendre P, Legendre LF. Numerical ecology, vol. 24. Elsevier; 2012.

Leiser CL, Hanson HA, Sawyer K, Steenblik J, Al-Dulaimi R, Madsen T, Gibbins K, Hotaling JM, Ibrahim YO, VanDerslice JA, et al. Acute effects of air pollutants on spontaneous pregnancy loss: a case-crossover study. Fertility and sterility. 2019; 111(2):341–347.

Lerchl A. Seasonality of sex ratio in Germany. Human reproduction (Oxford, England). 1998; 13(5):1401–1402.

Lobdell DT, Jagai JS, Rappazzo K, Messer LC. Data sources for an environmental quality index: availability, quality, and utility. American journal of public health. 2011; 101(S1):S277–S285.

Lynch R, Wasielewski H, Cronk L. Sexual conflict and the Trivers-Willard hypothesis: Females prefer daughters and males prefer sons. Scientific reports. 2018; 8(1):15463.

Melnikov V, Grech V. Seasonality of live birth sex ratio in south western Siberia, Russia, 1959–2001. Journal of Epidemiology & Community Health. 2003; 57(6):471–472.

Messer LC, Jagai JS, Rappazzo KM, Lobdell DT. Construction of an environmental quality index for public health research. Environmental Health. 2014; 13(1):39.

Murphy KP. Machine learning: a probabilistic perspective; 2012.

Orzack SH, Stubblefield JW, Akmaev VR, Colls P, Munné S, Scholl T, Steinsaltz D, Zuckerman JE. The human sex ratio from conception to birth. Proceedings of the National Academy of Sciences. 2015; 112(16):E2102–E2111.

R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2017, https://www.R-project.org/.

Samartsidis P, Seaman SR, Presanis AM, Hickman M, De Angelis D. Review of methods for assessing the causal effect of binary interventions from aggregate time-series observational data. arXiv preprint arXiv:180407683. 2018; .

Schwarz G. Estimating the dimension of a model. The annals of statistics. 1978; 6(2):461–464.

Scott SL. bsts: Bayesian Structural Time Series; 2019, https://CRAN.R-project.org/package=bsts, r package version 0.9.2.

Scott SL, Varian HR. Predicting the present with bayesian structural time series. Available at SSRN 2304426. 2013; .

Song S. Spending patterns of Chinese parents on children's backpacks support the Trivers-Willard hypothesis: Results based on transaction data from China's largest online retailer. Evolution and Human Behavior. 2018; 39(3):336–342.

Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006; 22(12):1540–1542.

Terrell ML, Hartnett KP, Marcus M. Can environmental or occupational hazards alter the sex ratio at birth? A systematic review. Emerging health threats journal. 2011; 4(1):7109.

Trivers RL, Willard DE. Natural selection of parental ability to vary the sex ratio of offspring. Science. 1973; 179(4068):90–92.

Zhang L, Liu W, Hou K, Lin J, Zhou C, Tong X, Wang Z, Wang Y, Jiang Y, Wang Z, et al. Air pollution-induced missed abortion risk for pregnancies. Nature sustainability. 2019; .

Figures and Tables

Exogenous Factor	Number of Studies	Sample Size
Dioxins (*Terrell et al., 2011*)	13	291
Polychlorinated biphenyls (PCBs) (*Terrell et al., 2011*)	9	98
1,2-Dibromo-3-chloropropane (DBCP) (*Terrell et al., 2011*)	2	29
Dichlorodiphenyltrichloroethane (DDT) (*Terrell et al., 2011*)	4	1623
Hexachlorobenzene (HCB) (*Terrell et al., 2011*)	2	262
Vinclozolin (*Terrell et al., 2011*)	1	95
Multiple pesticides (*Terrell et al., 2011*)	5	382
Lead (*Terrell et al., 2011*)	5	6566
Methylmercury (*Terrell et al., 2011*)	1	4808
Multiple metals (*Terrell et al., 2011*)	10	1015
Non-ionizing radiation (*Terrell et al., 2011*)	12	2926
Ionizing radiation (*Terrell et al., 2011*)	15	4959
Seasonality (*Lerchl, 1998*; *Melnikov and Grech, 2003*)	2	-
Ambient temperature (*Catalano et al., 2008*; *Helle et al., 2009*; *Fukuda et al., 2014*; *Dixson et al., 2011*)	4	-
Economic stress (*Catalano, 2003*)	1	-
Terrorist attacks (*James and Grech, 2017*)	2	-

Table 1. Exogenous factors reported in the literature to have an impact on the SRB (*Terrell et al., 2011*; *James and Grech, 2017*). A "-" indicates that sample sizes were not mentioned in the articles reporting or reviewing the corresponding results.

Factor name	effect
PCBs (air and water)	↑
DBCP (water)	–
Lead (land)	↓
Lead (air)	–
Aluminium (air)	↑
Chromium (air)	–
Chromium (water)	↑
Arsenic (land)	–
Arsenic (water)	↑
Cadmium (air and water)	–
Total mercury deposition	↑
Violent crime rate	–
Unemployed rate	–
Working out of county (long commute)	–

Table 2. Test results for factors selected from the literature reports (Table 1). A factor was included only if both its ΔIC and the coefficient of at least one of its septiles were statistically significant.

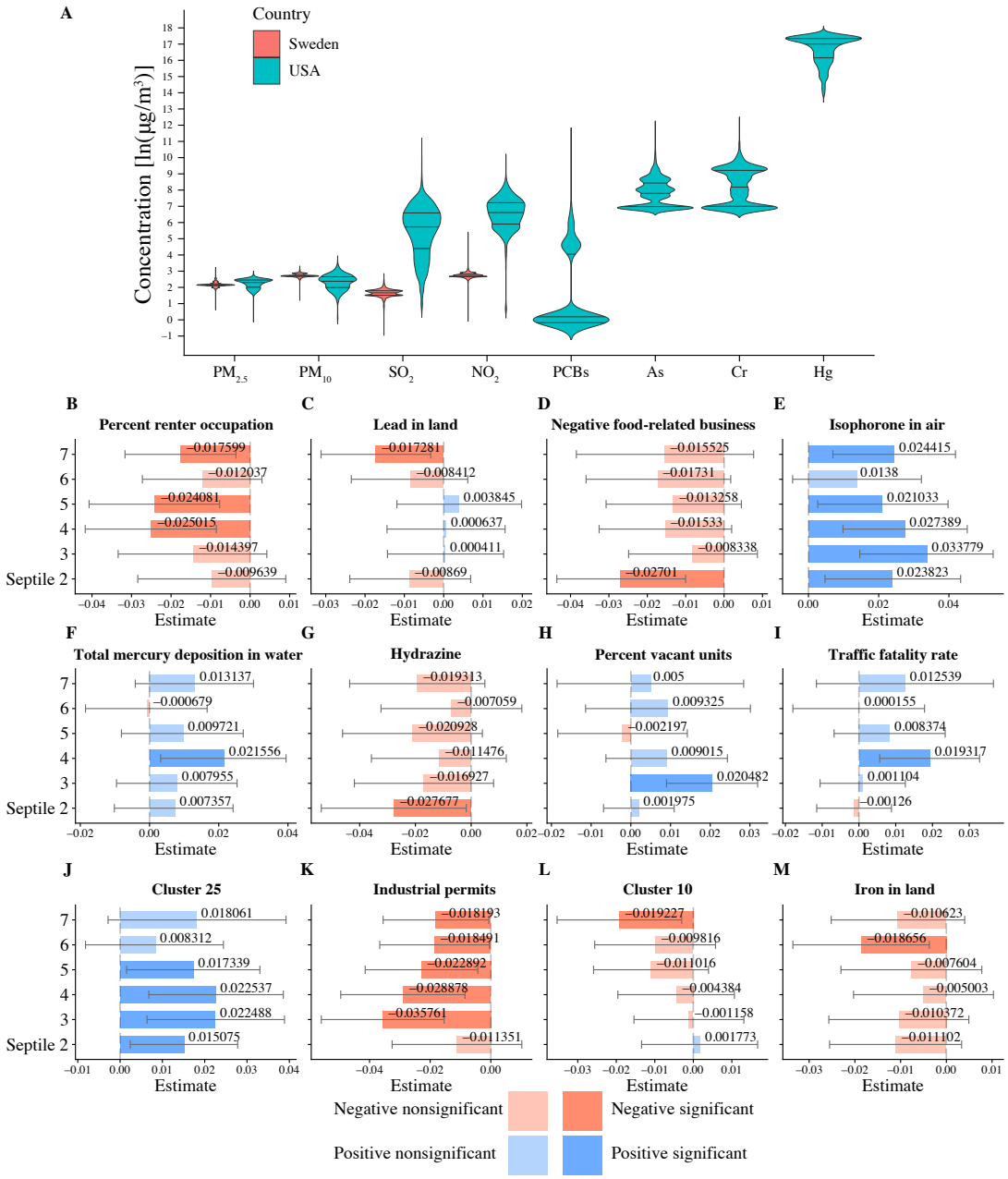

Figure 1. Airborne health-related substances and their association with the SRB. **A** Comparison of airborne pollutant concentrations across the US (cyan violin plots) and Sweden (pink violin plots). Only four air components, fine particulate matter ($PM_{2.5}$), coarse particulate matter (PM_{10}), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2) are measured in both countries. US counties appear to have higher mean pollution levels and are more variable in terms of pollution. **B-M**. A sample of 12 one-environmental factor logistic regression models that are most explanatory with respect to SRB. For each environmental factor, we partitioned counties into seven, equal-sized groups (septiles), ordered by levels of measurements, so that septile one corresponds to the lowest and septile seven to the highest concentration. Each plot shows bar plots of regression coefficients and 95% confidence intervals (error bar) of the second to the seventh septiles, with the first septile chosen as the reference level. The 12 models were ranked by the statistically significant factor's strength of association with at least one statistically significant coefficient by decreasing ΔIC; septiles whose coefficients are not significantly different from 0 at the 95% confidence level were plotted with a reduced alpha level. Blue bars represent positive coefficients, whereas red bars represent negative coefficients. "Negative food-related businesses" is a term used by the Environmental Protection Agency's Environmental Quality Index team and is explained as "businesses like fast-food restaurants, convenience stores, and pretzel trucks." "Percent vacant units" stands for "percent of vacant housing units." Substances contributing to cluster ten and cluster 25 are listed in Table 4. See Table S11 for more details regarding the factors' and clusters' identities.

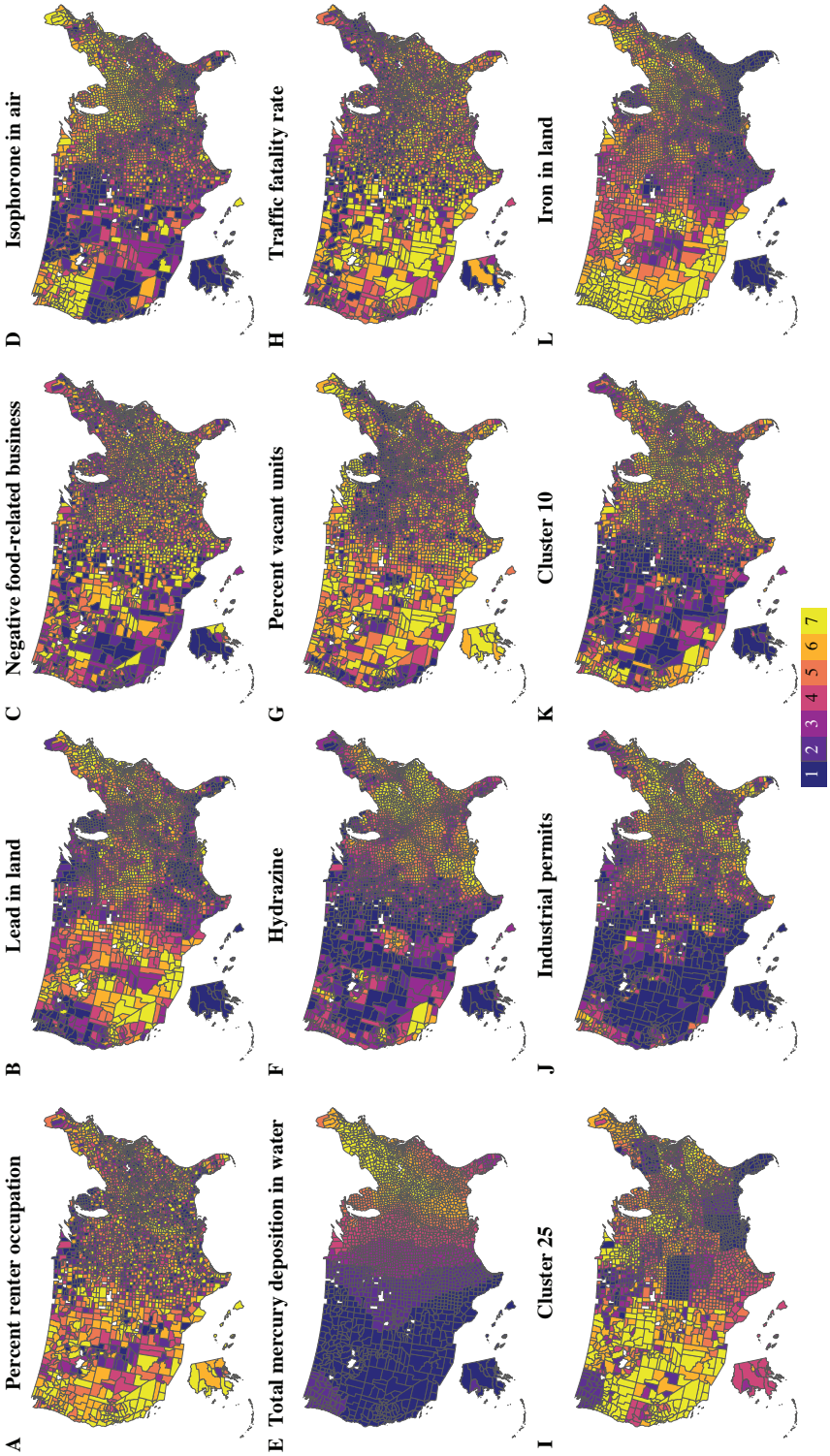

Figure 2. County-level geographical septile distribution for the first 12 statistically significant factors with at least one statistically significant coefficient ranked by decreasing ΔIC. The factors labelled **A–M** are the same as shown in Figure 1, Plates **B–M** and are ordered identically in both figures.

Factor name	effect
Iron	↓
Nitrate	↑
2-Nitropropane	↑
Carbon monoxide	↑
Bis-2-ethylhexyl phthalate	↓
Ethyl chloride	↑
Isophorone	↑
Hydrazine	↓
Phosphorus	↑
Quinonline	↓
Extreme drought	↑
Traffic fatality rate	↑
Industrial permits per $1000\,\mathrm{km}$ of stream	↓
Animal units	↓
Irrigation	↓
Negative food related businesses	↓
Renter occupation	↓
Vacant units	↑

Table 3. Test results for additional factors with statistically significant effects. A factor was included only if both its $\Delta\mathrm{IC}$ and the coefficient of at least one of its septiles were statistically significant.

Cluster number	factor
1	a_hcbd_ln,a_hccpd_ln
2	a_nitrobenzene_ln,a_dma_ln
3	a_2clacephen_ln,a_bromoform_ln
4	a_pnp_ln,a_toluene_ln
5	a_be_ln,a_se_ln
6	a_dmf_ln,a_edb_ln,a_edc_ln
7	a_teca_ln,a_procl2_ln,a_cl4c2_ln,a_vycl_ln,county_pop_2000
8	a_benzyl_cl_ln,a_me2so4_ln
9	mean_zn_ln,mean_cu_ln
10	mean_al_pct,mean_p_pct
11	numdays_close_activity_tot,numdays_cont_activity_tot
12	mean_as_ln,mean_se_ln
13	a_glycol_ethers_ln,a_etn_ln,a_vyac_ln
14	mean_na__pct_ln,mean_mg_pct_ln,mean_ca_pct_ln
15	a_cs_ln,a_edcl2_ln
16	a_ccl4,a_mtbe_ln
17	pct_harvest_acres,herbicides_ln,insecticides_ln
18	a_112tca_ln,a_ch3cn_ln
19	a_hcb_ln,a_pcp_ln,a_pcbs_ln
20	mg_ln_ave,k_ln_ave
21	pct_defoliate_acres_ln,pct_disease_acres_ln,pct_nematode_acres_ln
22	a_so2_mean_ln,a_no2_mean_ln,a_o3_mean_ln,so4_mean_ave
23	med_hh_value,med_hh_inc
24	rate_food_env_pos_log,rate_rec_env_log
25	ca_ln_ave,nh4_mean_ave
26	w_as_ln,w_ba_ln,w_cd_ln,w_cr_ln,w_cn_ln
	w_fl_ln,w_hg_ln,w_no3_ln,w_no2_ln,w_se_ln
	w_sb_ln,w_be_ln,w_ti_ln,w_endrin_ln
	w_lindane_ln,w_methoxychlor_ln,w_toxaphene_ln
	w_dalapon_ln,w_deha_ln,w_oxamyl_ln,w_simazine_ln
	w_dehp_ln,w_picloram_ln,w_dinoseb_ln
	w_hccpd_ln,w_carbofuran_ln,w_atrazine_ln
	w_alachlor_ln,w_heptachlor_ln,w_heptachlor_epox_ln
	w_24d_ln,w_silvex_ln,w_hcb_ln,w_benzoap_ln
	w_pcp_ln,w_124tcib_ln,w_pcb_ln,w_dbcp_ln
	w_edb_ln,w_xylenes_ln,w_chlordane_ln,w_dcm_ln
	w_odcb_ln,w_pdcb_ln,w_vcm_ln,w_11dce_ln
	w_t12dce_ln,w_edc_ln,w_111trichlorane_ln
	w_ccl4_ln,w_pdc_ln,w_trichlorene_ln,w_112tca_ln
	w_c2cl4_ln,w_cl1benz_ln,w_benzene_ln,w_toluene_ln
	w_ethylbenz_ln,w_stryene_ln,w_alpha_ln,w_dce_ln

Table 4. Clusters obtained by applying the Ward's method to the EQI dataset

Factor	explanation	ΔIC	SE
`pct_rent_occ`	percent renter occupation	54.2185	8.7760
`cluster_ward_9`	zinc (Zn) and copper (Cu) in air	52.3907	8.8256
`mean_pb_ln`	lead (Pb) in land	51.7339	9.6749
`farms_per_acre_ln`	farms per acre	51.6266	9.0878
`cluster_ward_15`	caesium (Cs) and ethylidene dichloride ($CHCl_2CHCl_2$) in air	51.5511	8.5623
`pct_mt_10units_log`	percent housing with $\geqslant 10$ units	51.3259	8.5483
`rate_food_env_neg`	negative food-related businesses (e.g. junk food)	51.3230	9.2907
`a_isophorone_ln`	isophorone in air	50.9430	10.3995

Table 5. List of the top 8 statistically significant environmental factors or clusters of factors with random effects at the county level from EQI by ΔIC. For the full table, see Table S3