Observable Variations in Human Sex Ratio at Birth

Yanan Long[1,2,3], Qi Chen[4], Henrik Larsson[4,5], Andrey Rzhetsky[2,3,6*],

1 Department of Chemistry, The University of Chicago, Chicago, IL, United States of America
2 Department of Medicine, The University of Chicago, Chicago, IL, United States of America
3 Institute of Genomics and Systems Biology, The University of Chicago, Chicago, IL, United States of America
4 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE-171 76, Sweden
5 School of Medical Sciences, Örebro University, Örebro, SE-701 82, Sweden
6 Department of Human Genetics and Committee on Quantitative Methods in Social, Behavioral, and Health Sciences, The University of Chicago, Chicago, IL, United States of America

* andrey.rzhetsky@uchicago.edu

Abstract

The human sex ratio at birth (SRB), defined as the ratio between the number of newborn boys to the total number of newborns, is typically slightly greater than 1/2 (more boys than girls) and tends to vary across different geographical regions and time periods. In this large-scale study, we sought to validate previously-reported associations and test new hypotheses using statistical analysis of two very large datasets incorporating electronic medical records (EMRs). One of the datasets represents over half ($\sim$ 150 million) of the US population for over 8 years (IBM Watson Health MarketScan insurance claims) while another covers the entire Swedish population ($\sim$ 9 million) for over 30 years (the Swedish National Patient Register). After testing more than 100 hypotheses, we showed that neither dataset supported models in which the SRB changed seasonally or in response to variations in ambient temperature. However, increased levels of a diverse array of air and water pollutants, were associated with lower SRBs, including increased levels of industrial and agricultural activity, which served as proxies for water pollution. Moreover, some exogenous factors generally considered to be environmental toxins turned out to induce higher SRBs. Finally, we identified new factors with signals for either higher or lower SRBs. In all cases, the effect sizes were modest but highly statistically significant owing to the large sizes of the two datasets. We suggest that while it was unlikely that the associations have arisen from sex-specific selection mechanisms, they are still useful for the purpose of public health surveillance if they can be corroborated by empirical evidences.

Author Summary

The human sex ratio at birth (SRB), usually slightly greater than 1/2, have been reported to vary in response to a wide array of exogenous factors. In the literature, many such factors have been posited to be associated with higher or lower SRBs, but the studies conducted so far have focused on no more than a few factors at a time and used far

smaller datasets, thus prone to generating spurious correlations. We performed a series of statistical tests on 2 large, country-wide health datasets representing the United States and Sweden to investigate associations between putative exogenous factors and the SRB, and were able to validate a set of previously-reported associations while also discovering new signals. We propose to interpret these results simply as public health indicators awaiting further empirical confirmation rather than as implicated in (adaptive) sexual selection mechanisms.

Introduction

Because human male gametes bearing X or Y chromosomes are equally frequent (being produced by meiosis symmetrically partitioning two sex chromosomes), and because ova bear only X chromosomes, one would expect a sex ratio at conception of exactly $\frac{1}{2}$ [1]. Indeed, a recent study using fluorescent *in situ* hybridization and array comparative genomic hybridization showed that the sex ratio at conception (SRC) was statistically indistinguishable from $\frac{1}{2}$ [2]. Nevertheless, the apparent sex ratio at birth (SRB), also known as the secondary sex ratio, has been documented to significantly deviate from $\frac{1}{2}$ under various circumstances, suggesting that a proportion of embryos are lost between conception and birth.

At least three processes may affect the observed SRB. First, female-embryo pregnancies may terminate early in development, driving the SRB up. It has been documented that these excess female-embryo losses tend to occur primarily during the first and early-second trimesters of pregnancy. Second, male-embryo deaths would drive the apparent SRB down. Male-embryo losses have indeed been observed to occur during the late-second and third trimesters [3]. Third, the SRB may be affected by maternal peri-conceptual hormonal levels [4, 5]. Past studies proposed that the SRB can fluctuate with time and may be driven by a number of environmental factors, such as chemical pollution, events exerting psychological stress on pregnant women (such as terrorist attacks and earthquakes), radiation, changes in weather, and even seasons of conception (Table 1).

Exogenous Factor	**Number of Studies**	**Sample Size**
Dioxins [6]	13	291
Polychlorinated biphenyls (PCBs) [6]	9	98
1,2-Dibromo-3-chloropropane (DBCP) [6]	2	29
Dichlorodiphenyltrichloroethane (DDT) [6]	4	1623
Hexachlorobenzene (HCB) [6]	2	262
Vinclozolin [6]	1	95
Multiple pesticides [6]	5	382
Lead [6]	5	6566
Methylmercury [6]	1	4808
Multiple metals [6]	10	1015
Non-ionizing radiation [6]	12	2926
Ionizing radiation [6]	15	4959
Seasonality [7, 8]	2	-
Ambient temperature [9–12]	4	-
Economic stress [13]	1	-
Terrorist attacks [14]	2	-

Table 1. Exogenous factors reported in the literature to have an impact on the SRB [6, 14]. A "-" indicates that sample sizes were not mentioned in the articles reporting or reviewing the corresponding results.

While there are multiple studies which have observed the positive associations between air pollution and spontaneous abortion [15, 16], most of those conclusions based on analyses of relatively small samples (Table 1), severely curtailing their statistical

power. In this study, we harnessed the power of 2 very large datasets: the MarketScan insurance claim data [17] in the United States (which records the health events of more than 150 million unique Americans, with more than 3 million unique newborns recorded between 2003 to 2011), and Sweden's birth registry data (covering the birth and health trajectories of over ~ 3 million newborns from 1983 to 2013) [18]. Our present study 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 [17] represents $104,565,671$ unique individuals and $3,134,062$ unique live births. The Swedish Health Registries [18] 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 an organization called Statistics Sweden. We used a subset of the MarketScan data that contained information on livebirths between 2003 to 2011 with county information encoded in Federal Information Processing Standards (FIPS) codes and a family link profile indicating the composition of the households 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) [19, 20].

Cluster Analysis

In order to simplify subsequent analyses, we first performed hierarchical clustering analysis on the Spearman's rank correlation coefficients matrix (ρ), using the Ward's method, which reduced the the EQI dataset's dimensionality. We then used the R-language [21] package `pvclust` [22] to minimize the total within-cluster variance [23]. 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 multilevel Bayesian logistic regression with random effects implemented in the R-language package `rstan` [24]. To facilitate model building, we used the R-language package `brms` [25] with default priors. Sampling was performed with the No-U-Turn sampler (NUTS) [26] with 500 warm-up steps and 1500 iteration steps with 28 Markov chains, of which the convergence was asseessed using the $\hat{R}$ statistic [27]. The model for the jth factor (predictor) is given as follows:

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

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, whenever applicable. The group-level effect was modeled for a single random effect by

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

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

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 [28]. 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 treated as baseline [29].

We applied logistic regression in two ways. First, to assess the effect of environmental factors, we regressed each of the individual factors' septiles against the SRB, with each sample point representing a county. Therefore, each septile, aside from the baseline, has a coefficient. Second, to test whether maternal 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 $\mathbf{x}_j = \mathbf{0}$, meaning, the model was comprised solely of the intercept). We computed the average difference in information criterion (ΔIC) and standard error (SE) for each factor obtained from leave-one-out (LOO) cross-validation [30], and used the Benjamini–Yekutieli method to adjust for multiple comparisons [31].

Univariate Time-Series Analysis

To assess the effect of one-off, stressful events on the SRB, we used two different time series techniques. First, we fitted seasonal univariate, autoregressive, integrated, moving average (sARIMA) models using the Box-Jenkins method [32], in conjunction with monthly (28-day periods) and weekly live birth data up to the event and then performed an out-of-sample prediction. An sARIMA model is given by

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

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"). Moreover, y_t indicates the observed univariate time series of interest, L is 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, chosen to be 4 in our study), and ϕ's, θ's, Φ's, and Θ's are model parameters to be estimated. We used the `auto.arima` function from the R-language package `forecast` [33,34] to fit the data, which performed a step-wise search on the (p, d, q, P, D, Q) hyperparameter space and compared different models by using the Bayesian Information Criterion (BIC) [35]. We confirmed the optimalx models' goodness-of-fit using the Breusch-Godfrey test on the residuals, which tested for the presence of autocorrelation, up to degree S [36–38].

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 [39]:

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

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

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 [39, 40]:

$$y_t = \mu_t + \tau_t + \varepsilon_t \tag{10}$$

$$\mu_t = \mu_t + \delta_{t-1} + u_t \tag{11}$$

$$\delta_t = \delta_{t-1} + v_t \tag{12}$$

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

Here, we define $\eta_t = \begin{bmatrix} u_t & v_t & w_t \end{bmatrix}$; Q_t is a t-invariant block diagonal matrix with diagonal elements σ_u^2, σ_v^2 and σ_w^2. Finally, we denote $\alpha_t = \begin{bmatrix} \mu_t & \delta_t & \tau_{t-S+2} & \cdots & \tau_t \end{bmatrix}$, 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` [41], which in turn relied on the R package `bsts` [42] 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 [43].

Results

We start by describing the negative results (i.e., a lack of a significant association), concordant across the two datasets. Our model selection rejected the whole spectrum of models that allow for periodic, annual SRB changes [7, 8]. For both US and Swedish datasets, the best-fitting model described the SRB as lacking seasonality throughout the year. Similarly, when we tested the claim that ambient temperatures during conception affect the SRB [9–12], we found that neither dataset supported this association. 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. Fig 1**A** shows a comparison of pollutant concentration distributions in both countries. The US environmental measurements dataset presented its own difficulty, as many pollutants appeared highly collinear in their spatial variation. To address this, we performed a cluster analysis on the environmental factors, subdividing them into 26 clusters (Table 2 and Fig S2). All pollutants within the same cluster were highly correlated, while the correlation between distinct clusters was much smaller, allowing for useful association inferences between SRB changes and environmental states.

Fig 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 4 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 partition counties into 7 equal-sized groups (septiles), ordered by levels of measurements, so that the first septile corresponds to the lowest and the highestnth septile 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. We rank the 12 models by the statistically significant factor's association strength with at least one statistically significant coefficient by decreasing ΔIC; septiles whose coefficients are not significantly different from 0 at the 95% confidence level have been 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 clusters 10 and 25 are listed in Table 2. See Table S11 for more details regarding the factors' and clusters' identities.

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 2. Pollutant clusters discovered by applying the Ward's method to the EQI raw measurements dataset.

Using the US dataset, we were able to validate the findings of a number of previous studies regarding the association between the SRB and exogenous factors (Table 3). Specifically, our data suggests that PCBs (polychlorinated biphenyls), aluminium (Al) in air, chromium (Cr) in water and total mercury (Mg) quantity drive the SRB up, while lead (Pb) in soil appears to be associated with a decreased SRB. Meanwhile, we have found no evidence for a number of previous reports, indicated with a dash in the second column in Table 3. We also established several new environmental associations that have not been reported before (Table 4, Fig 1 **B**–**M**, and Fig 2). Fig 1 **B**–**M** show 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).

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

Fig 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 Fig 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 km of stream	↓
Animal units	↓
Irrigation	↓
Negative food related businesses	↓
Renter occupation	↓
Vacant units	↑

Table 4. Test results for additional factors with statistically significant effects. We included a factor only if both its ΔIC and the coefficient of at least one of its septiles was statistically significant.

The geographic distribution of these pollutants varies remarkably, as seen in Fig 2. For example, lead in land (Fig 2**B**) appears to be enriched in the northeast, southwest, and mid-east US, but not in the south. Hydrazine (Fig 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 (Fig 2**E**) mostly affects eastern US states with the heaviest load in the northeastern states. It is this variability in the environmental distribution of various substances that allowed us to tease out these individual associations.

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 Figs 3C and 4C, cf. Tables S5c and S6c).

Fig 3. Time series plots and out-of-sample forecasts for SRB data grouped into 7-day periods and fitted with seasonal ARIMA models. The blue shade is the 95% confidence level. The observed SRBs for the first five months after the intervention are presented by red dots, whereas the observed SRBs for 7 to 9 months after the intervention are presented by purple dots. A: Hurricane Katrina, all states; B: Hurricane Katrina, Louisiana and Mississippi only; C: Virginia Tech shooting, all states; D: Virginia Tech shooting, adjacent states only.

Fig 4. Time series plots and out-of-sample forecasts for SRB data grouped into 7-day periods and fitted with state space models. The blue shade is the 95% confidence level. The observed SRBs for the first five months after the intervention are presented by red dots, whereas the observed SRBs for 7 to 9 months after the intervention are presented by purple dots. A: Hurricane Katrina, all states; B: Hurricane Katrina, Louisiana and Mississippi only; C: Virginia Tech shooting, all states; D: Virginia Tech shooting, adjacent states only.

Discussion

While SRB fluctuations in space and time are well-documented and non-controversial, there is a diverse range of competing theories striving to explain SRB changes in terms of mechanistic selective pressure [44]. The most frequently mentioned theory, is called the Trivers–Willard hypothesis (TWH), named after the researchers who proposed it [45]). The TWH postulates that, because the cost of rearing children is much higher for females than for males, in favourable, resource-rich environments, maleswould have more offspring than females, and vice versa in unfavourable conditions. Natural selection would then favour individuals with higher fitness, where fitness is equated to individuals' reproductive success (in this case, the number of offspring reaching reproductive age). According to the TWH, natural selection pushes the SRB up (more males) in favourable conditions, and down (more females) in unfavourable environment.

More explicitly stated, the TWH depends on the following three assumptions [5,45,46]:

Assumption A1. The condition of a mother during parental investment is correlated with the condition of her offspring; in other words, mothers in better conditions have offspring that will be in better conditions.

Assumption A2. The condition of the offspring persists after parental investment ends, and is positively correlated with the offspring's reproductive success.

Assumption A3. Males have larger variability in reproductive success than females and, as a result, they are more susceptible to sexual selection.

From these assumptions the TWH makes the following *deductive* inference on SRB variability:

Conclusion C1. The SRB varies such that females in favourable conditions have more male offspring, and in unfavourable conditions, more female offspring.

Assumption A3 is called Bateman's principle [47] (BP), and was suggested in a classic fruit fly genetics study on sexual selection. The original experimental results with *Drosophila melanogaster* indicated that males benefited more from multiple mating than females in terms of fitness. This asymmetry was thought to have originated in anisogamy, which means that a sperm is much smaller than an ovum and therefore requires less resources. Unfortunately, this result was never replicated (see [48, 49] for critiques of Bateman's methodology). Nevertheless, a modified version of BP, which generalizes anisogamy to parental investment, has enjoyed prominence among evolutionary biologists [50]. One of the critiques of BP claims that male cost of reproduction is in reality much higher than suggested by Bateman. This is because Bateman failed to account for the fact that males do not produce sperms stoichiometrically to match the number of female-produced ova. Instead, they produce semen, a mixture of a very large number of male gametes and accompanying secretions, rich in nutrients and other substances beneficial to reproduction [51]. Therefore, once the full range of investment patterns across life history (e.g. intrasex competition, secondary sexual characteristics, territorial defence) has been taken into account, it is unclear if reproductive investments of females exceed those of males [52, 53].

Faced with such criticisms, as well as an increasing amount of evidence from across the animal kingdom in which species did not conform to BP [54], supporters for BP have responded that sex differences ultimately originated from *historical* anisogamy [55, 56], and that there have also been subsequent ecological factors independent of anisogamy that drove sexual dimorphism, having to do with resource competition between the sexes, which may not result in stronger selection on males [57, 58]. Moreover, as a counter-challenge to the former point, Bateman's critics also refer to aggregate results in

favour of BP, including a phylogenic meta-analysis by Janicke et al., in which significant differences in reproductive success variances in species across the animal kingdom were found [59]. This reworking allowed for a potential remedy for BP, namely by generalizing it as follows [54, 60, 61]:

Assumption A3*. The sex with the larger reproductive success variance is more susceptible to sexual selection.

From this, the generalized version of the TWH follows:

Conclusion C1*. The SRB should vary such that females in favourable conditions have more offspring of that sex more susceptible to sexual selection, and in unfavourable conditions, more offspring of that sex less susceptible to sexual selection.

This version of BP is consistent with "sex-role reversals" observed in many species, in which females exhibit larger susceptibility to sexual selection. In addition, it allows for sufficient flexibility such that the identity of "the sex more susceptible to sexual selection" may be influenced by exogenous conditions [57]. Candidates for the identity of that sex include higher variance in number of (adult) offspring and higher variance in parental investments [5, 62]. Nevertheless, under this revised framework, for sexually dimorphic selection patterns to develop and persist, as opposed to randomly fluctuating across time [63], one inevitably has to invoke the sexual cascade hypothesis; namely that a small initial difference (e.g. anisogamy) in sex-related phenotype will "snowball" into larger, persistent patterns through hereditary feedback loops [64–66]. Such cascading has also featured in the above-mentioned meta-analysis discussion regarding high-level explanatory patterns among animal species [59], bracketing all differences in sex-related traits into one-dimensional sexual selection [54]. There is a plethora of other competing theories, e.g. [63] and [61], which predict largely stochastic variations of sex-related phenotypes, emphasizing the role social and ecological factors have played in shaping *plastic* sex-roles [54, 67, 68]. Even if the last point may still be somewhat contentious [65], BP and (by extension) the TWH are, at the very least, not the only game in town when it comes to explaining and predicting patterns related to sexual selection: male and female phenotypes of a given species in a given environment are most likely the results of a large number of exogenous factors without any single one of them being particularly dominant [69, p. 177].

One key ramification of the above analysis is that the TWH cannot provide a comprehensive account of the range of exogenous factors associated with SRB variation under the kind of circumstances present in our study. Further, the empirical success of the TWH is mixed, with only 50% of studies confirming it, and around 20% of studies producing statistically significant results in the opposite direction [46], which is consistent with our finding that many different pollutants might be assumed to be "bad" for mothers (e.g. pollutants, traffic fatality rates, junk food) had associations with SRB in *opposite* directions. The scepticism against the applicability of the TWH in contemporary human societies is further strengthened by two recent population studies in Sweden with large sample sizes (4.7 and 5.7 million live births, respectively), which found no SRB heritability [70, 71]. In particular, Zietsch et al. have demonstrated that there exists neither within-individual SRB auto-correlation (contra Assumption A1) nor similarity in the SRB for children of siblings (contra Assumption A2). They also concluded that within-family SRB was associated with the final family size, suggesting that SRB variations may have been the result of SRB-*aware* family planning [70]. Taken together, such evidence also places other adaptive (i.e. via heritable sexual selection) theories explaining SRB variations, such as adaptive versions of hormonal hypothesis [62], maternal dominance hypothesis [72, 73] and the Bruce effect [74] in the same predicament. Thus, our results are better interpreted as supporting the intrinsic randomness of the

SRB and/or the dependence of the SRB on *non-adapative* (e.g. socio-cultural [70, 75, Ch. 12]), causal factors, possibly including those common to both changes in the SRB and associated exogenous factors.

By way of conclusion, we note that the literature includes substantial reports on the relationship between the SRB and public health [3], and we would like to consider the question of whether the SRB can be used as an indicator for public health events and, if so, whether the relationship between the SRB and certain diseases reveals causal relationships. As the preceding discussion demonstrates, even if the existence of *adaptive* causal relationships between environmental factors and the SRB may be unlikely (contra [76]), *associations* – including the ones presented in this work – may be used as signals for (adverse) public health conditions, as long as they are established experimentally. To this end, we reiterate that there *are* agreements between the associations established in our work and those in the literature [6, 14], and that our results do support the non-monotonous, dose-response profiles frequently reported in the literature [77] (Table 3). Therefore, future research programmes might instead focus on exploring and validating the associations between SRB and environmental factors that reliably *predict* adverse public health effects for certain subpopulations [71] using large datasets with covariates sampled frequently across considerable spatio-temporal ranges [70]. Another interesting direction would be to determine the potentially non-adaptive physiological mechanisms.

Limitations

Unlike some of the recent studies [78], we did not have access to the sex of stillbirths, which would have enabled us to probe negative selection *in utero* against frail males [3]. When quantifying pollutants in the US, we used the EPA air quality raw data, which was an average of measurements taken over a short period of time, rather than over years or decades, which would have enabled long-term and causal analyses. Neither did it include information for individual exposures to those factors, which might render a straightforward interpretation of our results subject to ecological fallacies. Finally, the subjects in our US study were commercially-insured and had medical claims, which likely came from a different probability distribution to the general population in the US.

Supporting Information

S1 Appendix Additional results, figures and tables

Acknowledgments

We are grateful to E. Gannon and M. Rzhetsky for comments on earlier versions of this manuscript.

References

1. Fisher RA. The genetical theory of natural selection. The Claredon Press; 1930.
2. Orzack SH, Stubblefield JW, Akmaev VR, Colls P, Munné S, Scholl T, et al. The human sex ratio from conception to birth. Proceedings of the National Academy of Sciences. 2015;112(16):E2102–E2111.
3. Bruckner TA, Catalano R. Selection in utero and population health: theory and typology of research. SSM-population health. 2018;5:101–113.
4. Grant VJ, Chamley LW. Can mammalian mothers influence the sex of their offspring peri-conceptually? Reproduction. 2010;140(3):425–433.
5. James WH. Evolution and the variation of mammalian sex ratios at birth: reflections on Trivers and Willard (1973). Journal of theoretical biology. 2013;334:141–148.
6. 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.
7. Lerchl A. Seasonality of sex ratio in Germany. Human reproduction (Oxford, England). 1998;13(5):1401–1402.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. Catalano RA. Sex ratios in the two Germanies: a test of the economic stress hypothesis. Human Reproduction. 2003;18(9):1972–1975.
14. 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.
15. Leiser CL, Hanson HA, Sawyer K, Steenblik J, Al-Dulaimi R, Madsen T, et al. Acute effects of air pollutants on spontaneous pregnancy loss: a case-crossover study. Fertility and sterility. 2019;111(2):341–347.
16. Zhang L, Liu W, Hou K, Lin J, Zhou C, Tong X, et al. Air pollution-induced missed abortion risk for pregnancies. Nature sustainability. 2019;.

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

18. 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.

19. 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.

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

21. R Core Team. R: A Language and Environment for Statistical Computing; 2017. Available from: `https://www.R-project.org/`.

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

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

24. Stan Development Team. RStan: the R interface to Stan; 2020. Available from: `http://mc-stan.org/`.

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

26. 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.

27. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC. Bayesian analysis. 2021;1(1):1–28.

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

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

30. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and computing. 2017;27(5):1413–1432.

31. 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.

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

33. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. Journal of Statistical Software. 2008;26(3):1–22.

34. Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, et al.. forecast: Forecasting functions for time series and linear models; 2019. Available from: `http://pkg.robjhyndman.com/forecast`.

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

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

37. 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.

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

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

40. Murphy KP. Machine learning: a probabilistic perspective. The MIT Press; 2012.

41. 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.

42. Scott SL. bsts: Bayesian Structural Time Series; 2019. Available from: `https://CRAN.R-project.org/package=bsts`.

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

44. Douhard M. Offspring sex ratio in mammals and the Trivers-Willard hypothesis: In pursuit of unambiguous evidence. BioEssays. 2017;39(9):1700043.

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

46. James WH. Possible constraints on adaptive variation in sex ratio at birth in humans and other primates. Journal of Theoretical Biology. 2006;238(2):383–394.

47. Bateman AJ. Intra-sexual selection in Drosophila. Heredity. 1948;2(3):349–368.

48. Hoquet T, Bridges WC, Gowaty PA. Bateman's Data: Inconsistent with "Bateman's Principles". Ecology and Evolution. 2020;10(19):10325–10342.

49. Gowaty PA, Kim YK, Anderson WW. No evidence of sexual selection in a repetition of Bateman's classic study of Drosophila melanogaster. Proceedings of the National Academy of Sciences. 2012;109(29):11740–11745.

50. Trivers R. Parental investment and sexual selection. Sexual Selection & the Descent of Man, Aldine de Gruyter, New York. 1972; p. 136–179.

51. Hrdy SB. Empathy, polyandry, and the myth of the "coy" female. In: Bleier R, editor. Feminist Approaches to Science. Pergamon Press; 1986. p. 119–146.

52. Hubbard R. The politics of women's biology. Rutgers University Press; 1990.

53. Gerlach N, McGlothlin J, Parker P, Ketterson E. Reinterpreting Bateman gradients: multiple mating and selection in both sexes of a songbird species. Behavioral Ecology. 2012;23(5):1078–1088.

54. Tang-Martínez Z. Rethinking Bateman's principles: challenging persistent myths of sexually reluctant females and promiscuous males. The Journal of Sex Research. 2016;53(4-5):532–559.

55. Schärer L, Rowe L, Arnqvist G. Anisogamy, chance and the evolution of sex roles. Trends in Ecology & Evolution. 2012;27(5):260–264.

56. Kokko H, Booksmythe I, Jennions MD. Causality and sex roles: prejudice against patterns? A reply to Ah-King. Trends in ecology & evolution. 2013;28(1):2–4.

57. Morimoto J. Bateman (1948): was it all wrong? A comment on Hoquet (2020). Animal Behaviour. 2020;168:e1–e4.

58. De Lisle SP. Understanding the evolution of ecological sex differences: Integrating character displacement and the Darwin-Bateman paradigm. Evolution Letters. 2019;3(5):434–447.

59. Janicke T, Häderer IK, Lajeunesse MJ, Anthes N. Darwinian sex roles confirmed across the animal kingdom. Science advances. 2016;2(2):e1500983.

60. Arnold SJ, Duvall D. Animal mating systems: a synthesis based on selection theory. The American Naturalist. 1994;143(2):317–348.

61. Hoquet T. Bateman's principles: why biology needs history and philosophy. Animal Behaviour. 2020;168:e5–e9.

62. James WH. Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. The Journal of endocrinology. 2008;198(1):3–15.

63. Gowaty PA, Hubbell SP. Reproductive decisions under ecological constraints: it's about time. Proceedings of the National Academy of Sciences. 2009;106(Supplement 1):10017–10024.

64. Parker GA. The sexual cascade and the rise of pre-ejaculatory (Darwinian) sexual selection, sex roles, and sexual conflict. Cold Spring Harbor perspectives in biology. 2014;6(10):a017509.

65. Parker GA, Pizzari T. Sexual selection: the logical imperative. In: Current perspectives on sexual selection. Springer; 2015. p. 119–163.

66. Fromhage L, Jennions MD. Coevolution of parental investment and sexually selected traits drives sex-role divergence. Nature communications. 2016;7(1):1–11.

67. Gowaty PA. Standing on Darwin's shoulders: the nature of selection hypotheses. In: Current perspectives on sexual selection. Springer; 2015. p. 103–118.

68. Roughgarden J. Sexual selection: Is anything left? Current Perspectives on Sexual Selection. 2015; p. 85–102.

69. Fine C. Testosterone Rex: Myths of sex, science, and society. WW Norton & Company; 2017.

70. Zietsch BP, Walum H, Lichtenstein P, Verweij KJ, Kuja-Halkola R. No genetic contribution to variation in human offspring sex ratio: a total population study of 4.7 million births. Proceedings of the Royal Society B. 2020;287(1921):20192849.

71. Catalano R, Casey JA, Bruckner TA. A test of oscillation in the human secondary sex ratio. Evolution, Medicine, and Public Health. 2020;2020(1):225–233.

72. Grant VJ. The maternal dominance hypothesis: questioning Trivers and Willard. Evolutionary Psychology. 2003;1(1):147470490300100106.

73. Grant VJ. Could maternal testosterone levels govern mammalian sex ratio deviations? Journal of Theoretical Biology. 2007;246(4):708–719.

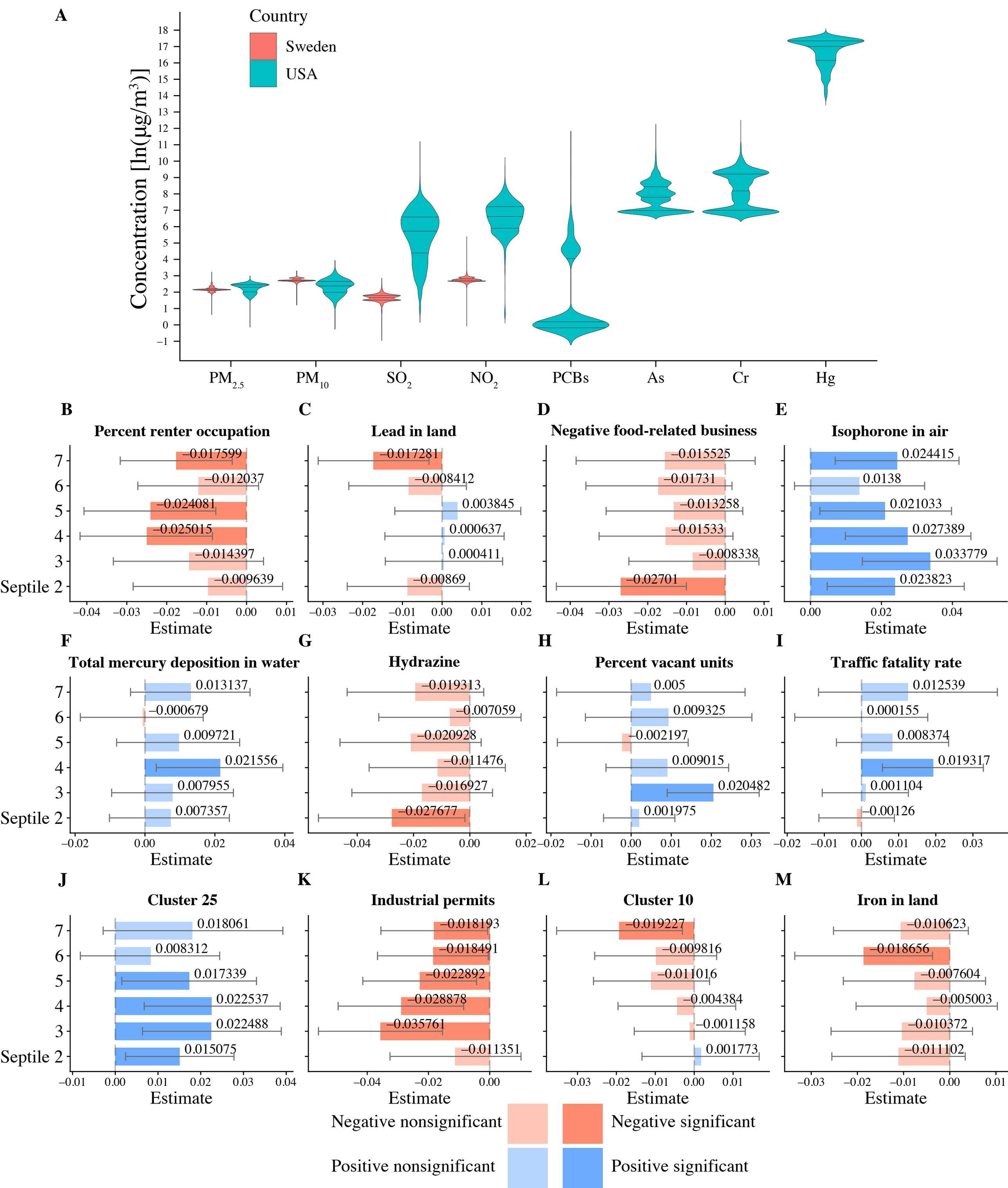
74. Catalano R, Gemmill A, Casey J, Karasek D, Stewart H, Saxton K. Separating the Bruce and Trivers-Willard effects in theory and in human data. American Journal of Human Biology. 2018;30(2):e23074.

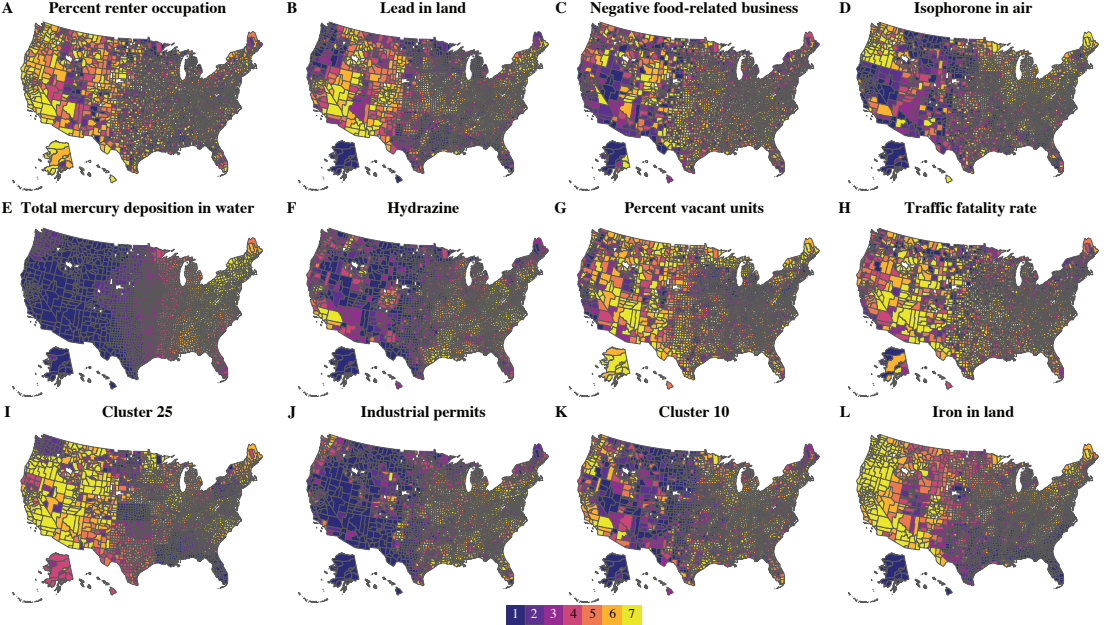
75. Dupré J. Processes of life: Essays in the philosophy of biology. Oxford University Press; 2012.

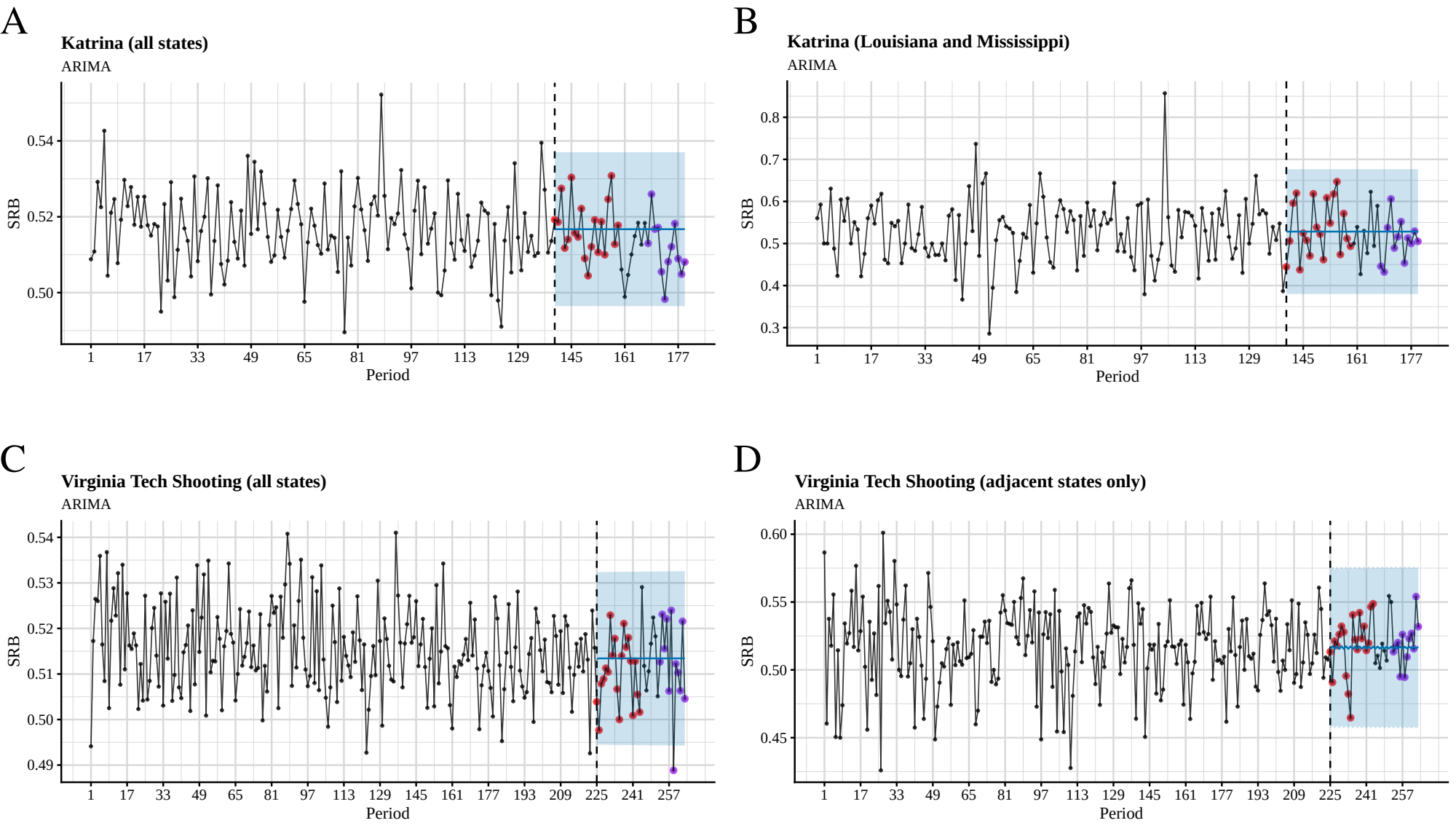
76. James WH, Grech V. Can sex ratios at birth be used in the assessment of public health, and in the identification of causes of selected pathologies? Early human development. 2018;118:15–21.

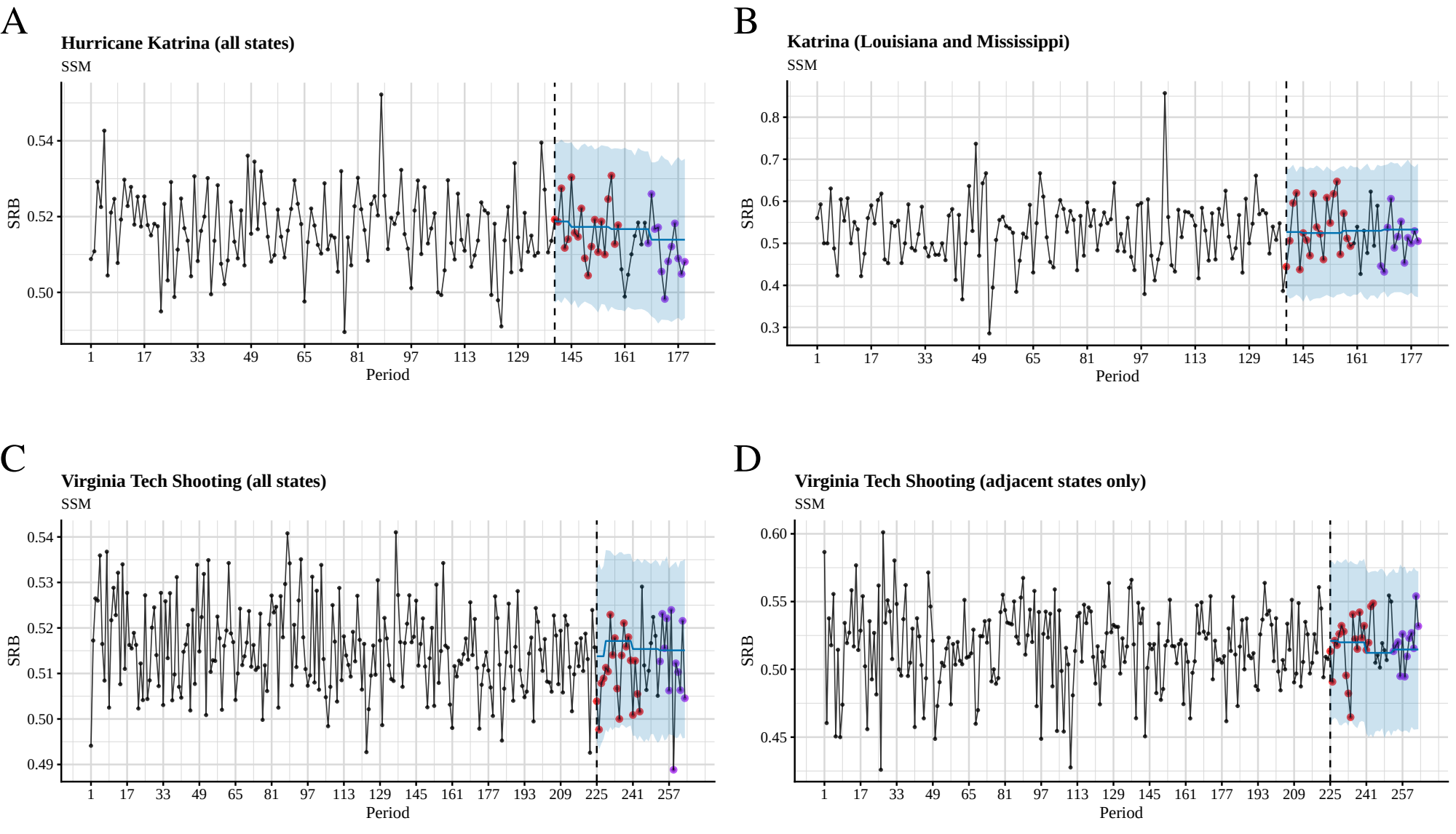
77. Pavic D. A review of environmental and occupational toxins in relation to sex ratio at birth. Early human development. 2020;141:104873.

78. Catalano R, Bruckner T, Gould J, Eskenazi B, Anderson E. Sex ratios in California following the terrorist attacks of September 11, 2001. Human Reproduction. 2005;20(5):1221–1227.


A
Country
Sweden
USA
Concentration [ln(μg/m3)]
PM2.5
PM10
SO2
NO2
PCBs
As
Cr
Hg
B
Percent renter occupation
−0.017599
−0.012037
−0.024081
−0.025015
−0.014397
−0.009639
C
Lead in land
−0.017281
−0.008412
0.003845
0.000637
0.000411
−0.00869
D
Negative food-related business
−0.015525
−0.01731
−0.013258
−0.01533
−0.008338
−0.02701
E
Isophorone in air
0.024415
0.0138
0.021033
0.027389
0.033779
0.023823
F
Total mercury deposition in water
0.013137
−0.000679
0.009721
0.021556
0.007955
0.007357
G
Hydrazine
−0.019313
−0.007059
−0.020928
−0.011476
−0.016927
−0.027677
H
Percent vacant units
0.005
0.009325
−0.002197
0.009015
0.020482
0.001975
I
Traffic fatality rate
0.012539
0.000155
0.008374
0.019317
0.001104
−0.00126
J
Cluster 25
0.018061
0.008312
0.017339
0.022537
0.022488
0.015075
K
Industrial permits
−0.018193
−0.018491
−0.022892
−0.028878
−0.035761
−0.011351
L
Cluster 10
−0.019227
−0.009816
−0.011016
−0.004384
−0.001158
0.001773
M
Iron in land
−0.010623
−0.018656
−0.007604
−0.005003
−0.010372
−0.011102
Septile
Estimate
Negative nonsignificant
Negative significant
Positive nonsignificant
Positive significant


A
Percent renter occupation
B
Lead in land
C
Negative food-related business
D
Isophorone in air
E
Total mercury deposition in water
F
Hydrazine
G
Percent vacant units
H
Traffic fatality rate
I
Cluster 25
J
Industrial permits
K
Cluster 10
L
Iron in land
1
2
3
4
5
6
7


A
Katrina (all states)
ARIMA
SRB
Period
B
Katrina (Louisiana and Mississippi)
ARIMA
SRB
Period
C
Virginia Tech Shooting (all states)
ARIMA
SRB
Period
D
Virginia Tech Shooting (adjacent states only)
ARIMA
SRB
Period


A
Hurricane Katrina (all states)
SSM
SRB
0.54
0.52
0.50
1 17 33 49 65 81 97 113 129 145 161 177
Period
B
Katrina (Louisiana and Mississippi)
SSM
SRB
0.8
0.7
0.6
0.5
0.4
0.3
1 17 33 49 65 81 97 113 129 145 161 177
Period
C
Virginia Tech Shooting (all states)
SSM
SRB
0.54
0.53
0.52
0.51
0.50
0.49
1 17 33 49 65 81 97 113 129 145 161 177 193 209 225 241 257
Period
D
Virginia Tech Shooting (adjacent states only)
SSM
SRB
0.60
0.55
0.50
0.45
1 17 33 49 65 81 97 113 129 145 161 177 193 209 225 241 257
Period