THE UNIVERSITY OF CHICAGO

PROBING THE HEALTH OUTCOMES IN EARLY CHILDHOOD USING BAYESIAN DATA ANALYSIS ON BIG DATA

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY
YANAN LONG

CHICAGO, ILLINOIS
AUGUST 2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS

$\mathcal{D}$	Data
f	Probability density function (PDF)
F	Cumulative distribution function (CDF)
h	Hazard function
H	Cumulative hazard function
Γ	Gamma distribution/function
ℓ	Logarithmic probability density function/likelihood
$\mathcal{N}$	(Multivariate) normal distribution
S	Survival function

ACKNOWLEDGMENTS

First, I would like to express my heartfelt gratitude for my advisor Andrey Rzhetsky, who has afforded me immense patience and support throughout my time in his research group. I am especially grateful for introducing me to the exciting fields of biomedical data science and Bayesian statistics when I was a junior graduate student seeking to switch fields. Thank you also to members of the Rzhetsky group: Atif Khan, Chao Zhang, Hanxin Zhang, Rachel Melamed, Gengjie Jia, Bohdan Khomtchouk and Chengjian Shi. In particular, I am impressed by Atif's meticulous attention to detail, and it has been a great pleasure to work with him.

I am also grateful for my dissertation committee members, Profs Raymond Moellering, Aaron Dinner and Michael Rust, whose questions and comments during and after the defence have provided distinct perspectives to my work and helped me improved the exposition of some of the theoretical background.

Furthermore, I would like to thank my collaborators from the Risi Kondor group, Horace Pan, Truong Son Hy and Erik Thiede — I have learnt a great deal from their formidable expertise in geometric deep learning. On this note I would also like to thank Prof Eric Jonas for introducing me to the field of machine learning and encouraging me to pursue further studies.

Dr Vera Dragisich has been a phenomenal presence in the Chemistry department — I am thankful for her constant efforts to make the PhD experience better for everyone.

To all my fellow organizers at GSU: thank you for the camaraderie outside of academics and I am anxiously anticipating our eventual victory.

To my new friends at Queer in AI: thank you for accepting me into the community and I am proud of what we have achieved.

Finally, I am grateful for my parents' unwavering support through the ups and downs of graduate school: I could not have done this without you.

ABSTRACT

Pregnancy is a fascinating biological process in which the best theory and evidence suggest that the mother and newborn form two divergent processes which are themselves integrated into the surrounding environment. Thus, the early life health profile of a newborn depends closely on that of their mother as well as environmental factors during pregnancy. While many previous studies have provided evidence for such associations, they either tend to have small sample sizes and hence limited statistical power, or only have studied a few of the factors in isolation. In this study, we systematically probed the associations by means of principled Bayesian data analysis on a very large commercial insurance claims dataset based in the United States. After a brief exposition of the prerequisite of the statistical models employed in our study, we first identified newborns in the dataset with potential mothers using appropriate diagnostic codes, which served as the foundation of our subsequent work. Using this matched newborn–mother cohort, we first re-examined more than a hundred previously reported associations between the sex ratio at birth, also known as the secondary sex ratio, with environmental, socioeconomic and other stress factors. By showing that these association did not form discernible patterns that popular adaptive theories of sexual selection predict, we provided further evidence against these theories and call for a reformulation of this particular area of evolutionary theory for humans. To demonstrate further use of the matched cohort, we performed a large cohort study on the associations between early-life neurodevelopmental disorders and maternal immune activation, use of anti-infective prescription and various adverse birth conditions. Echoing recent results suggesting a common aetiology of early life immune system diseases and neurodevelopmental disorders as disruptions to modules of the microbiota–gut–brain axis, our results demonstrated that many risk factors that may be effects of dysbiosis were indeed associated with elevated risks of neurodevelopmental disorders. Finally, we included a preliminary study on using Morgan fingerprints as an efficient method of attaining state-of-the-art predictive performance on a commonly used drug–drug interactions dataset.

CHAPTER 1

INTRODUCTION

Pregnancy is a fascinating biological process whereby new mammals organisms are generated. Traditionally assumed to be either containing the conceptus (embryo/foetus) like a bun in the oven or subsuming it as a proper part, the gravida (i.e. the entity being pregnant with the conceptus) is much more closely integrated metabolically and immunologically than previously thought (Kingma, 2019; Finn, 2023). At stake here is the concept of biological individuality, namely when and where a (new) organism begin to exist (Guay and Pradeu, 2016a). Latest developments in theoretical biology have marked a shift away from a strict substance view on individuality, according to which organisms have well-defined boundaries, akin to objects studied in e.g. many branches physics. Instead, mounting empirical evidences from across the different kingdoms suggest that seemingly discrete individuals criss-crosses other such individuals at multiple length- and timescales (cf. Dupré (2020)). As such, pregnancy may be most appropriately viewed as two (or more, in the case of multiple gestation) diverging processes that, while the pregnancy lasts, are highly intertwined (Meincke, 2022). More generally, organisms are increasing being treated as *processes* rather than things in theoretical biology, giving rise to the new doctrine of processualism Nicholson and Dupré (2018)

In a similar vein, earlier immunology has painted a picture in which a biological organism, equipped with an immune system, fend off external pathogens, as posited by the self–non-self theory (Pradeu, 2020, §§2.1 & 3.1). Under this theoretical framework, the host ("self") learns to preserve itself and defend against foreign elements ("non-self") such as infectious agents and xenografts. Despite intuitive clarity and empirical success, the self–non-self theory has run into a multitude of theoretical issues. For instance, autoimmunity poses a curious challenging on how to unambiguously delineate the boundary between the self and the non-self. More importantly, there exist a large number (in humans, as many as the genetically "self" cells) of genetically-foreign microorganisms that are not only not eliminated by the host's immune system, but play an active role in the proper

functioning and regulation of the immune system. The latter scenario has prompted many to consider complex multicellular organisms (e.g. animals and plants) as "holobionts": the "self" along with its symbionts (Gilbert, 2014; Van de Guchte et al., 2018). This symbiosis is maintained by the so-called "negotiated surveillance", where the host is equipped with an *evolved* mechanism to incorporate potentially self-replicating parts into its wholes, a process that upends classical understanding of such complex organisms as "anatomical, developmental, physiological, immunological, genetic, or evolutionary" individuals (Gilbert et al., 2012). The new field of ecological evolutionary developmental biology (Eco-Evo-Devo) (Gilbert et al., 2015; Sultan, 2021) takes up the insight and further interrogates the way in which the co-construction shapes.

In the case of humans specifically, the role that the host–microbiota equilibrium play in health and diseases has been increasingly recognized. *Dysbiosis*, the disruption of the equilibrium which results generally in reduced taxonomic diversity and accumulation of large taxa, have been reported to cause a wide range of diseases (Proal et al., 2017; Brown and Clarke, 2017; Walker, 2017; Sokol et al., 2017). In autoimmune and inflammatory diseases, for example, multiple immunological pathways involving various types of T helper cells have been identified (Dehner et al., 2019), whereas in infectious diseases, besides activation of pathways, the symbiotic microbiota also promote the secretion of antimicrobial proteins and peptides (APPs) by leukocytes (Brown and Clarke, 2017).

While the microbiota are located primarily in the human gut, pioneering studies in immunology have revealed the enormous extent to which the immune and the nervous systems are intimately related, giving rise the emergent subfield of neuroimmunology. This used to be a somewhat surprising discovery, given that the brain had generally been regarded as immune-privileged: the absence of immune cells in the brain, which is separated from the immune system by the blood-brain-barrier (BBB). However, this has been demonstrated false on both accounts: not only is the BBB not entirely impermeable, the brain also has its own immune system which interacts with the rest of the immune system at multiple levels (Pradeu, 2020, §5.3), including that of the intestine. As such, a considerable body of recent work has focused on the the microbiota–gut–brain (MGB) axis, in

particular the way in which dysbiosis may trigger neuropsychiatric disorders (Cryan et al., 2019; Jacobson et al., 2021).

To add a further twist, recent work has highlighted the effects of health conditions of the human mothers on their offspring. Of particular importance is the transmission of the newborn microbiota mediated by that of the mother. The former consensus view among biologists of the foetus growing in a sterile uterus and acquiring colonies of bacteria while passing through the birth canal is all but refuted by the finding that intrauterine microbiota are very similar to that of the vaginal tract (Gilbert, 2014). More intriguing perhaps is the discovery that the the meonatal gut microbiota was dissimilar from both the maternal gut and vaginal microbiota, but resembelled that of the placenta and the oral cavity. Taken together, we have sketched portrait of the human immune system that develops early on in pregnancy from a foetus–mother–microbiota community, with the foetus progressively acquiring microbiota and diverging from the maternal immune system as it develops into a full-fledged holobiont Takeshita (2022).

Having introduced the biological problems we are about to investigate, we now present the methodology for tackling these problems. Recent years have seen a growing interest in applying *Bayesianism* to scientific reasoning. A key motivations is that the researchers' belief in a scientific theory should be allowed to come in degrees, calibrated by data gathered from experiments, rather than an all-or-nothing affair (Lin, 2022). Although precise details might vary for different authors, the vast majority of them accepts the following commitments (Bird, 2017; Sprenger and Hartmann, 2019, Theme: p. 28):

- Probabilism: the degrees of belief (credences) of researchers are represented by probability distributions that follow the axioms of probability;

- Conditionalization: such credences shall be updated based on Bayes theorem, that is, given new data/evidence $\mathcal{D}$, the new credence (posterior) of a model/hypothesis $\mathcal{M}$ is equal to the old credence (prior) conditioned on that evidence $\mathbb{P}(\mathcal{M}|\mathcal{D}) = \frac{P(\mathcal{M})\,P(\mathcal{D}|\mathcal{M})}{P(\mathcal{D})}$ with $P(\mathcal{D}) > 0$.

The model $\mathcal{M}$ in the above is usually taken to be parameterized by $\boldsymbol{\theta}$ that constitutes the locus of statistical inference. $\mathcal{M}$ is considered generative, i.e. a joint probability distribution of the parameters and the data.

While usually construed as an inductive mode of reasoning, Bayesianism, especially in practice such as in this dissertation, is probably better seen as a form of inference to the best explanation (IBE), also known as *abductive reasoning* (or simply *abduction*), which contrasts with deduction where the result of an inference must follow from the premises from which it is inferred (though some prominent practitioners of Bayesianism characterize it as deductive, cf. (Gelman and Shalizi, 2013)). As a non-necessary form of inference, abduction is often distinguished from the similar *induction*, which is based *solely* on statistics (Douven, 2021, 2022). The distinction here is that rather focusing exclusively on posterior probability, other *explanatory* factors also play an instrumental role in proper scientific reasoning: these may include predictive accuracy, simplicity, elegance, and other pragmatic considerations, reflected in the way *model selection* (better termed *model comparison*) is carried out (Sprenger and Hartmann, 2019, Variations 7 & 11). There is another sense in which proper statistical modelling in the sciences is not simply a matter of maximizing the (posterior) probability, but rather, according to the statistician George Box, that:

> the statistician knows, for example, that in nature there never was a normal distribution, there never was a straight line, yet with normal and linear assumptions, known to be false, he [*sic*] can often derive results which match, to a useful approximation, those found in the real world. (Box, 1976, §2.5, p. 792)

The above is usually expressed in the pithy aphorism "all models are wrong, but some are useful." As Box's example shows, in scientific modelling one inevitably makes uses of *idealization*, which means that in Bayesian modelling one's credence about scientific theories/hypotheses are conditioned on the idealized model, which is called *suppositional analysis* in (Sprenger and Hartmann, 2019, §12.2). Formally, one assumes that the model $\mathcal{M}$ is given by $\mathcal{M} = (\mathcal{S}, \mathbb{P})$, where $\mathcal{S}$ is the sample space and $\mathbb{P}$ the set of probability distributions on $\mathcal{S}$. Now the choice of $\mathcal{M}$, of

course, depends on the modelling aims of the researcher. Then, all evaluation of probability within this idealization is implicitly conditioned on $\mathcal{M}$. The consequence is that (Sprenger and Hartmann, 2019, §12.3, p. 322):

> we should not read off our actual degrees of belief from a Bayesian model; instead, the model informs our degrees of belief and our predictions by showing what they would be under reasonable idealizing assumption,

and that

> [the] soundness [of Bayesian inference] depends on whether *the overall model is well chosen or inadequate*. (Emphasis added)

1.1 Outline

The remainder of the dissertation is organized as follows:

- In Chapter 2, we briefly review the theoretical background of the statistical analysis that we perform in subsequent chapters, including measure-theoretic probability theory, regression analysis using the generalized linear model, the Bayesian workflow, and Markov Chain Monte-Carlo.

- In Chapter 3, we show how to match newborns in the MarketScan dataset to mothers and provide summary statistics for the various cohorts used in subsequent studies.

- In Chapter 4, we study the human sex ratio at birth (SRB), defined as the ratio between the number of newborn boys to the total number of newborns, which is typically slightly greater than $1/2$ (more boys than girls) and tends to vary across different geographical regions and time periods. We evaluate previously-reported associations as well as new hypotheses using statistical analysis of two very large datasets incorporating electronic medical records

(EMRs): the IBM MarketScan (cf. Ch. 3) and the Swedish National Patient Register (*Patientregistret*). After testing more than 100 associations, 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 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 further empirical evidences.

- In Chapter 5, continuing the theme of mother–newborn link, we jointly probe association between disorders affecting the nervous system and the immune system using again the MarketScan dataset. Although it is only relatively recently recognized how deeply the two systems are intertwined and how the proper functioning of them depend on each other — as evidenced by the research programme of the microbiota–gut–brain (MGB) axis — our study represents the only one that makes use of a unified cohort to study early-onset neurodevelopmental disorders (NDDs) and challenges to the health of the immune system of both the newborn and the mother. Bayesian cross validation showed that while the count regression models for immune system phenotypes failed the stress test of posterior predictive check, the logistic regression models for neurodevelopmental disorders, covering autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD) and learning difficulties, passed them. These models yielded results broadly in congruence with existing literature, predicting lower risks of girls vs boys, and higher risks of newborn delivered via Caesarean section and very premature births. For ADHD only, we also found that both maternal and

newborn uses of prescription antiinfectives, as well as higher levels of $PM_{2.5}$ pollution, were associated with higher risks. Moreover, we also observed interesting results regarding time of birth: there was quasi-monotonic trend in risks of NDD diagnosis as a function of year of birth, and, for ADHD we detected non-linear effects of time of birth in year on the risks. While direct causal claims were not warranted under our study design, our large sample size and unified approach lent support to our suggestion that further studies should test the role of immune activation in NDDs.

- In Chapter 6, we demonstrate that for the problem of drug–drug interaction (DDI) prediction a simple neural networks using Morgan fingerprints of drugs outperformed these more complicated GNN models while spending only a small fraction of the time in training. Furthermore, to improve training, we curated and made available a novel dataset with negative drug–drug interaction examples derived from a very large electronic health records dataset. By contrast, contemporary machine learning studies tackle the drug–drug interaction forecast problem by featurizing drugs using graph neural networks (GNNs). This automated featurization allows to avoid laborious handcrafting chemical features.

- In Chapter 7 we summarize the main results obtained in the previous chapters and conclude this dissertation.

CHAPTER 2

BACKGROUND

In this chapter, we first give precise of definitions of relevant mathematical objects required for the statistical models that we will employ later.

2.1 Probability Theory

2.1.1 Preliminaries

Definition 2.1 (Collection of Sets). A **topology** $\mathcal{T}$ is a collection (i.e. set) of subsets of a set X such that:

1. $\varnothing, X \in \mathcal{A}$;
2. $A_1, A_2 \in \mathcal{T} \implies A_1 \bigcap A_2 \in \mathcal{T}$;
3. $\underbrace{A_1, A_2, \ldots}_{\text{countably many}} \in \mathcal{T} \implies \bigcup_i A_i \in \mathcal{T}$.

An **algebra** $\mathcal{A}$ over X is a collection of subsets such that:

1. $\varnothing, X \in \mathcal{A}$;
2. $A \in \mathcal{A} \implies A^{\complement} \in \mathcal{A}$;
3. $A_1, \ldots, A_n \in \mathcal{A} \implies \bigcup_{i=1}^{n} A_i \in \mathcal{A} \wedge \bigcap_{i=1}^{n} A_i \in \mathcal{A}$.

Moreover, $\mathcal{A}$ becomes a σ-**algebra** if the last condition is changed into

4. $\underbrace{A_1, A_2, \ldots}_{\text{countably many}} \in \mathcal{A} \implies \bigcup_i A_i \in \mathcal{A} \wedge \bigcap_i A_i \in \mathcal{A}$.

We say that a topology or an algebra is **generated by** a set or another collection of subsets called the generator if the generated set is the smallest topology or algebra containing the generator. Finally, the **Borel σ-algebra**, which unifies topology and σ-algebra, is the σ-algebra generated by a topology. □

Definition 2.2 (Measure). Let X be a set and Σ a σ-algebra over X. A measure μ is a function from Σ to $\overline{\mathbb{R}}$ satisfying:

- $\mu(\varnothing) = 0$.
- $\forall A \in \Sigma\colon \mu(A) \geqslant 0$.
- Given a family $\left\{A_j\right\}_{j=1}^{\infty}$ of disjoint subsets of Σ, $\mu\left(\bigcup_j A_j\right) = \sum_j \mu\left(A_j\right)$.

Furthermore, a measure is σ-finite if there exists $\left\{A_j\right\}_j$ such that $X = \bigcup_j A_j$ and $\forall j\colon \mu(A_j) < \infty$.

Definition 2.3 (Measure and Probability Space). A **measurable space** is a 2-tuple (X, Σ), where X is a non-empty set and Σ a σ-algebra on X. By extension, a **measure space** is a 3-tuple (X, Σ, μ) where $\mu\colon \Sigma \mapsto E$, where $E \subseteq \overline{\mathbb{R}}$. A probability space is a measure space when $E = [0, 1]$ and $\mu(X) = 1$, and is often written instead as $(\Omega, \mathcal{F}, \mathbb{P})$.

A common interpretation of the above formalism for probability spaces is this: the elements of the *sample space* X are the possible outcomes of some observation or experiment, whereas the *event space* Σ contains the events of interest, which are themselves sets of outcomes, and finally the *probability measure* assigns each event to a real number in $[0, 1]$. In statistics, the most commonly seen measures are the counting measure and the Lebesgue measure:

Definition 2.4 (Counting Measure). Given a set X and a σ-algebra over X, the **counting measure** is defined as:

$$\mu(A) = |A| \tag{2.1}$$

where $|\cdot|$ the cardinality of a set.

Definition 2.5 (Lebesgue Measure). Let $\mathcal{B}$ be the Borel σ-algebra on $\mathbb{R}$ generated by finite open intervals. The measure on $(\mathbb{R}, \mathcal{B})$ given by

$$\mu([a, b]) = b - a \tag{2.2}$$

for $-\infty < a \leqslant b < \infty$ is the **Lebesgue measure**.

A measure space $(X, \mathcal{F}, \mathbb{P})$, as defined above, can be a bit too abstract and arbitrary, and so we often need to map it to another space that is easier to work with. In order to do this we shall first introduce the concept of a measurable function

Definition 2.6 (Measurable Function). Let (X, Σ) and $(Y, \mathcal{G})$ be measurable spaces. A function $f\colon X \mapsto Y$ is said to be **measurable** if for any $G \in \mathcal{G}$ we have $f^{-1}(G) \in \mathcal{F}$.

2.1.2 Random Variables

With the definition of probability measure in hand, we can now proceed to the central object in probability theory — random variables.

Definition 2.7 (Random Variable). Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable space $(M, \mathcal{M})$, an **M-valued random variable (RV)** $X\colon \Omega \mapsto M$ is a measurable function. Then, $\mathbb{P}$ induces a **pushforward measure** $\mathbb{P}_X$ on $(M, \mathcal{M})$ given by

$$\mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right) \tag{2.3}$$

for every $B \in \mathcal{M}$. This induced measure is called the **distribution** of X, and we say that X has distribution $\mathbb{P}_X$, or $X \sim \mathbb{P}_X$.

Definition 2.8 (Distribution). In applications, however, we are mostly interested in real-valued random variables, in which case $M = \mathbb{R}^k$ for some positive integer k and $\mathcal{M}$ is the Borel σ-algebra on

$\mathbb{R}^k$. In this case, the (joint) **cumulative distribution function (CDF)** is defined as

$$F(x_1, \ldots, x_k) = \mathbb{P}_X((-\infty, x_1], \ldots, (-\infty, x_k]), \tag{2.4}$$

The **complementary cumulative distribution function (CCDF)** is defined as

$$\overline{F}(x_1, \ldots, x_k) = \mathbb{P}_X((x_1, \infty), \ldots, (x_k, \infty)). \tag{2.5}$$

In survival analysis, CCDF is known as the **survival function** and denoted as $S(\cdot)$. Finally, the marginal CDF is defined as

$$F_j(x) = \lim_{x_{j'} \to \infty, j' \in [1,k]-\{j\}} F(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_k). \tag{2.6}$$

Proposition 2.9. Some useful properties of CDF are (cf. (Shao, 2003, Exercise 10, p. 75)):

1. $x_k \leqslant x'_k \implies F(x_1, \ldots, x_{k-1}, x_k) \leqslant F(x_1, \ldots, x_{k-1}, x'_k)$.
2. For all $j \in [1, k] \bigcap \mathbb{Z}$, we have $\lim_{x_j \to -\infty} F(x_1, \ldots, x_k) = 0$ and $\lim_{x_j \to \infty} F(x_1, \ldots, x_k) = 1$.

Definition 2.10 (Independence). Let F be the joint CDF of RVs $X_1, \ldots, X_k$ and F_j the marginal CDF of X_j for $j \in [1, k] \bigcap \mathbb{Z}$. Then the RVs are said to be **independent** if

$$F(x_1, \ldots, x_k) = \prod_{j=1}^{k} F(x_j). \tag{2.7}$$

In application, we are often in the position of assessing the average behaviour of a random variable over all values on which it is defined. Hence, we need some notion of summation:

Definition 2.11 (Lebesgue Integration). Let $(X, \mathcal{F}, \mu)$ be a measure space. A function $s \colon X \mapsto \mathbb{R}$

is called a **simple function** if it can be expressed as

$$s(\cdot) = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j}(\cdot), \tag{2.8}$$

where the A_j's are measurable subsets of X and $\mathbb{1}_E(\cdot)$ is the **indicator function** on the set E, that is

$$\mathbb{1}_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}. \tag{2.9}$$

Given a simple function s, the **Lebesgue integral** w.r.t the measure μ is given by

$$\int s \, d\mu = \sum_{j=1}^{n} \alpha_j \mu\left(A_j\right). \tag{2.10}$$

Now let f be a measurable function $X \to \overline{\mathbb{R}}_+$. Then Lebesgue integral of f w.r.t. the measure μ is defined as the supremum of all , given by

$$\int f \, d\mu = \sup\left\{ \int s \, d\mu \middle| 0 \leqslant s \leqslant f \right\}. \tag{2.11}$$

Though not required for probability measures, we can generalize the definition to $\overline{\mathbb{R}}$-valued functions by simply partitioning the integrand into positive and negative parts:

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu, \tag{2.12}$$

where

$$\begin{cases} f_+(\cdot) & = \max(f(\cdot), 0) \\ f_-(\cdot) & = \max(-f(\cdot), 0) \end{cases}. \tag{2.13}$$

We can pursue a different kind of generalization by means of integrating w.r.t. a different measure called the **Lebesgue–Stieltjes** measure, defined as

$$\mu_{\mathrm{LS}} = F(b) - F(a) \qquad \text{for all } a, b \in \mathbb{R} \text{ with } a < b. \tag{2.14}$$

We can then define the **Lebesgue–Stieltjes integral** w.r.t. F as the Lebesgue integral w.r.t. the measure μ_{LS}, written as $\int f\, dF$. □

A special class of integral is the expectation or expected value[1].

Definition 2.12 (Expected Value). Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Then the **expected value** of X, written as $\mathbb{E}[X]$, is given by

$$\int X\, d\mathbb{P}, \tag{2.15}$$

and if F_X is the CDF of $\mathbb{P}$ on $\mathbb{R}^k$, we also write

$$\mathbb{E}[X] = \int dF(x). \tag{2.16}$$

We now present a central result in probability theory which enables easy computation using less well-known measures. To do so, we first observe that if μ is a measure on (X, Σ), then

$$\lambda(A) = \int_A f\, d\mu, \tag{2.17}$$

for a non-negative Borel function is itself a measure on (X, Σ), with $\mu(A) = 0 \implies \lambda(A) = 0$. If we can find the requisite f we will be able to convert direct computation using λ with integration over μ. The following result gives the sufficient condition for the existence of such f:

Theorem 2.13 (Radon–Nikodym Theorem). Let μ and λ be σ-finite measures on some measurable

1. Not expectation value, as is usually referred to in the physics literature, since it is a pleonasm.

space (X, Σ). We say that λ is **absolutely continuous** w.r.t. μ if $\mu(A) = 0 \implies \lambda(A) = 0$, written as $\lambda \ll \mu$.

Suppose that $\lambda \ll \mu$. Then there exists a non-negative Borel function f such that Equation 2.17 holds, which is unique in the sense that if there is another function g that also satisfies Equation 2.17, then f and g are identical except on a set of measure 0 w.r.t. μ. The last condition can be written as $f = g$ μ-a.e., which stands for **almost everywhere**.

The function f is the called the **Radon–Nikodym derivative** of λ w.r.t. μ, denoted $\frac{d\lambda}{d\mu}$. For a probability space, if $\int f \, d\mu = 1$ and $f \geqslant 0$ μ-a.e., then λ defined in 2.17 is the **probability density function** (PDF) w.r.t. μ. When μ is the counting measure, we call f a discrete PDF, and when μ is the Lebesgue measure, we call f a continuous PDF. □

We shall present examples of both discrete and continuous PDFs below in §2.1.3, but before we proceed let us consider an illuminating example: a PDF that is neither discrete nor continuous, namely the PDF of a continuous distribution subject to right-censoring.

Example 2.14 (Continuous CDF with Constant Right-Censoring). Let T be an RV on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ whose CDF F_T has a Lebesgue PDF. Let c be a constant such that $F_T(c) < 1$. We now define another RV $X = \min(T, c)$. This type of setup is typical of in experiments with right-censoring, where the observed value of the RV of interest T is only up to a certain upper bound c.

Since $F_Y(t)$ is discontinuous at c, Y does not have a Lebesgue PDF. It does not have a discrete PDF either, as it is not absolutely continuous w.r.t. the counting measure on $(\infty, c]$. However, it is absolutely continuous w.r.t. the measure $\mu + \delta_c$, where $\delta_c(A) = \mathbb{1}_A(c)$. □

An addition use of Equation 2.17 is in the definition of conditional expectation, which is of the bedrock of Bayesian statistics.

Definition 2.15 (Conditional Expectation). Let X be an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$,

that is

$$\int |X| \, d\mu < \infty. \tag{2.18}$$

Then the **conditional expectation** of X given a sub-σ-algebra $\mathcal{A}$ of $\mathcal{F}$ is an a.s.-unique random variable $\mathbb{E}[X|\mathcal{A}]$ such that for all $A \in \mathcal{A}$

$$\int_A \mathbb{E}[X|\mathcal{A}] \, d\mathbb{P} = \int_A X \, d\mathbb{P}. \tag{2.19}$$

Moreover, for any event $B \in \mathcal{F}$, the **conditional probability** of B given $\mathcal{A}$ is given by

$$\mathbb{P}(B|\mathcal{A}) = \mathbb{E}[\mathbb{1}_B|\mathcal{A}]. \tag{2.20}$$

The notion of the conditional expectation of X given another RV Y is also well-defined:

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)]. \quad \square \tag{2.21}$$

Finally, the **conditional distribution** of a RV given another one can be obtained from the conditional expectation.

Theorem 2.16 (Conditional distribution). Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{A}$ a sub-σ-algebra of $\mathcal{F}$. Then, by Theorem 2.13, there exists a function $\mathbb{P}(B, \omega)$ on $\mathcal{B}(\Omega) \times \Omega$ such that

- for any $B \in \mathcal{B}(\Omega)$, we have $\mathbb{P}(B, \omega) = \mathbb{P}\left(X^{-1}(B)\middle|\mathcal{A}\right)$ a.s.;
- for any $\omega \in \Omega$, the function $\mathbb{P}(\cdot, \omega)$ is a probability measure on $(\mathbb{R}^n, \mathcal{B})$;

Moreover, if Y is a measurable function from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\Lambda, \mathcal{G})$, then there exists a function $\mathbb{P}_{X|Y}(B|y)$ such that

- for any $B \in \mathcal{B}(\Omega)$, we have $\mathbb{P}_{X|Y}(B|y) = \mathbb{P}\left(X^{-1}(B)\middle|Y = y\right)$ a.s.;

- for any $y \in \Lambda$, the function $\mathbb{P}_{X|Y}(\cdot|y)$ is a probability measure on $(\mathbb{R}^n, \mathcal{B})$. □

2.1.3 Probability Distributions

Most distributions presented below are *parametric*, although in Bayesian statistics *non-parametric* and *semi-parametric* statistical models are also widely used.

Definition 2.17 (Parametric Family (Shao, 2003, §2.1.2)). Let $\mathbb{P}_{\boldsymbol{\theta}}$ be a set of probability measures on a measurable space (X, Σ) indexed by a **parameter** $\boldsymbol{\theta} \in \Theta$. $\mathbb{P}_{\boldsymbol{\theta}}$ is said to be a **parametric family** if $\Theta \subseteq \mathbb{R}^d$ for some positive integer d and each element of $\mathbb{P}_{\boldsymbol{\theta}}$ is known whenever $\boldsymbol{\theta}$ is known. The set Θ is called the **parameter space** and d its **dimension**. A parametric family is said to be **identifiable** if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2 \implies \mathbb{P}_{\boldsymbol{\theta}_1} \neq \mathbb{P}_{\boldsymbol{\theta}_2}$.

Perhaps the most widely used family of probability distribution in statistics is the exponential family (expfam).

Definition 2.18 (Exponential Family). A parametric family $\{\mathbb{P}_{\boldsymbol{\theta}}\}$ of CDFs with PDFs w.r.t. a σ-finite measure μ on a probability space $(\Sigma, \mathcal{F})$ is called an exponential family if

$$\frac{d\mathbb{P}_{\boldsymbol{\theta}}}{d\mu}(\omega) = h(\omega) \exp\left[\boldsymbol{\eta}(\boldsymbol{\theta}) \cdot \mathbf{T}(\omega) - A(\boldsymbol{\theta})\right]. \tag{2.22}$$

Here, $h(\omega)$ is called the base measure, $\boldsymbol{\eta}$ the natural parameter, $A(\boldsymbol{\theta})$ the log-partition function and $\mathbf{T}(\omega)$ the sufficient statistic.

The notion of a statistic being "sufficient" introduced in Definition 2.18 for estimating a parameter w.r.t. a sample if, roughly speaking, no other statistic can provide any additional information regarding the value of the parameter. A formal definition is as follows.

Definition 2.19 (Sufficient Statistics). Let X be a sample for some parametric family $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$. Then a statistic $T(X)$ is said to be **sufficient** if the conditional distribution of X given T does not depend on θ. □

The notion of a sufficient statistic in the context of the exponential family turns out to be handy in improving the computational efficiency of some statistical models (c.f. Theorem 2.21 below).

2.2 Regression Analysis

In statistics, we often pose our question in the form of modelling the relationship between dependent variables (the outcome or response) and independent variables (the predictors) using regression analysis. A common assumption to make is that the dependent variables can be expressed as a function of some linear combination of the independent variables. The model that encodes this assumption is the generalized linear model (GLM), which is a very popular class of regression models which extends the ordinary linear model (OLM). In this case, the response variable is allowed to be dependent on the predictors via a link function. While GLMs reduce to the OLM when the link function is the identity function and the outcome follows the normal distribution, one crucial difference between the general case and the special case is that the likelihood equation may not have a closed-form expression.

Definition 2.20 (Generalized Linear Model (Agresti, 2015, Ch. 1)). The generalized linear model consists of three components:

1. **random component**: this is the outcome variable $\boldsymbol{y}$;
2. **linear predictor**: this is the design matrix X along with the model parameters $\boldsymbol{\beta}$;
3. **link function**: this is the function that directly maps from the *mean* of the random component to the linear predictor

$$\mathbb{E}[\boldsymbol{y}] = g^{-1}\left(X^\top \boldsymbol{\beta}\right). \tag{2.23}$$

Finally, we require that the $\boldsymbol{y}$ be distributed as an exponential family. □

Below are a few useful properties of the GLM.

Theorem 2.21 (Properties of GLM). 1. The sufficient statistics for the model parameters $\boldsymbol{\beta}$ are

$$T(\boldsymbol{\beta}) = \sum_{n=1}^{N} y_n x_{nm} \quad m \in [1, p] \bigcap \mathbb{Z} \tag{2.24}$$

$$= X^\top \mathbf{y}. \tag{2.25}$$

Thus, we can obtain the same inference for $\boldsymbol{\beta}$ if we group entries X and add up the corresponding values in $\mathbf{y}$.

2. If the natural parameter of the exponential family part is given by $\boldsymbol{\eta}(\theta) = \boldsymbol{\eta}\left(X^\top \boldsymbol{\beta}\right)$, then the regression coefficients $\boldsymbol{\beta}$ have **conjugate priors** — such that the posterior and prior belong to the same parametric family — given by

$$\mathbb{P}_\pi(\boldsymbol{\beta} | \boldsymbol{\chi}, \nu) = f(\boldsymbol{\chi}, \nu) \exp\left[\boldsymbol{\chi}^\top \boldsymbol{\beta} - \nu A(\boldsymbol{\beta})\right], \tag{2.26}$$

where $\boldsymbol{\chi}$ and ν are hyperparameters. □

2.3 Bayesian Workflow

In this section, we provide a general Bayesian workflow (Gelman et al., 2020) that will be employed in the subsequent chapters. This workflow, which consists of the three steps of model building, inference and model checking, enables us to handle uncertainty that arises in various aspects of the analytical process in a disciplined manner. Figure 2.1 gives a high-level overview of how one might go about with the workflow, bearing in mind that not every step in the workflow will necessarily be carried out in a given study.

Before beginning, we assume that a dataset is already available and fixed, and while this is the case for our studies, it may not be so in general: for example, in drug trials the data-collection process is ongoing until the conclusion of the trial. In either case, the first step in the workflow is to pick (an) initial model(s), which is usually drawn from prior research. Ideally, these models should

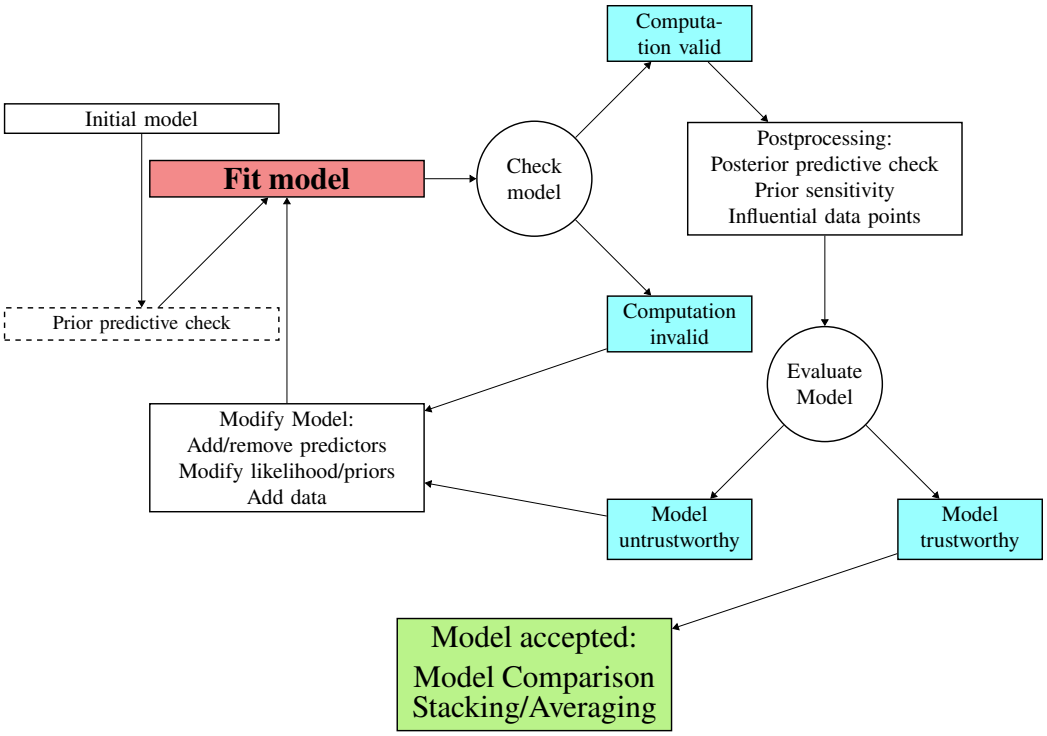


Figure 2.1: An overview of the Bayesian workflow, simplified from (Gelman et al., 2020)

capture the main effect that the researcher aims to model and are relative easy to fit. After specifying the initial models, we can optionally perform the prior predictive check, which uses draws from the prior predictive distribution, that is the distribution of the unknown but observable y:

$$\mathbb{P}(y) = \int \mathbb{P}(\theta)\mathbb{P}(y|\theta)\, d\theta. \tag{2.27}$$

In effect, this is doing fake data simulation to generate possible observed data in order to assess how plausible the model is given the data.

The next step is fitting the model(s) to the actual observed data using MCMC (cf. §2.4) or alternatively some approximation method such as integrated nested Laplace approximation (INLA) (Rue et al., 2017). Since we will be using a variant of Hamiltonian Monte Carlo (HMC) in all our studies, we shall assume that a whole gamut of model checking methods are available. After model fits terminate either successfully or due to out-of-time (OOT) or out-of-memory (OOM) errors, we

proceed to the validation of the computation obtained. In the case of OOT or OOM, one could try to reduce either the model by removing some predictors or using simplified forms of the predictors, or use a subset of the data. On the other hand, if the computation terminated, it can be validated using the diagnostic tools that come with HMC (cf. §2.5). If any abnormalities are detected, the model components may need to be modified (e.g. likelihood/prior or the functional forms of the predictors) until the results are valid.

If the computation turns out to be valid, we can now post-process the model fit to evaluate its trustworthiness. Possible actions include prior sensitivity analysis, posterior predictive check (PPC), cross validation (CV) and examination of the influences of particular data points, the last two of which is usually done using leave-one-out cross validation (LOO-CV) (cf. §2.6). PPC is done in an analogous way to Equation 2.27 but now with all the observed data

$$\mathbb{P}\left(\tilde{y}|\, y\right) = \int \mathbb{P}\left(\theta|\, y\right) \mathbb{P}\left(\tilde{y}|\, \theta\right)\, d\theta. \tag{2.28}$$

Similar to the previous step, in the event that the model fails to pass some set of the stress test, we shall modify the model components until the results are trustworthy.

2.4 Markov Chain Monte-Carlo

In this section, we provide a brief exposition of the theory of MCMC, and specifically the No-U-Turn sampler (NUTS) variant of the HMC sampler (Hoffman et al., 2014), which is the main workhorse that powers all Bayesian data analysis performed in our studies. The essence of MCMC is to create a Markov chain of which the stationary distribution is the target of the simulation. A Markov chain, moreover, is nothing more than a "memoryless" stochastic process.

Definition 2.22 (Markov Chain). A **transition kernel** π is a function on $X \times \mathcal{B}(X)$, where $\mathcal{B}(\cdot)$ is the Borel σ-algebra on a set, that satisfies the following conditions:

- $\pi(x, \cdot)$ is a probability measure for all $x \in X$;

- $\pi(\cdot, A)$ is measurable for all $A \in \mathbb{B}(X)$.

Intuitively, the transition kernel encodes the conditional probability density of the states given an initial state x, that is

$$\mathbb{P}(A|x) = \int_A \pi(x, x')\, dx'. \tag{2.29}$$

Given a transition kernel π, we can define the **Markov chain** induced by it as a sequence of random varaiables $\{X_n\}$ such that for any t the conditional distribution of X_t on all previous entires in the sequence is equal to the conditional distribution of X_t on just the last element, X_{t-1}. More specifically

$$\mathbb{P}(X_t \in A|X_0 = x_0, \cdots, X_{t-1} = x_{t-1}) = \mathbb{P}(X_t \in A|X_{t-1} = x_{t-1}) \tag{2.30}$$

$$= \int_A \pi(x_k, dx). \quad \square \tag{2.31}$$

Definition 2.23 (MCMC). A Markov Chain Monte-Carlo (MCMC) method for generating a sample from some target f is any method producing an ergodic Markov chain with stationary distribution f. $\square$

We are now ready to introduce the classical Metropolis–Hastings algorithm which serves as the foundation for many improved MCMC methods, including HMC and NUTS.

Definition 2.24 (Metropolis–Hastings). The **Metropolis–Hastings algorithm** is an acceptance/rejection sampling method that converges to the given target. The algorithm is as follows:

1. For $t = 0$, draw a sample of the starting point θ^0 from a starting distribution $f_0(\theta)$.
2. For $t > 0$ draw a sample θ^* from the proposal distribution $J_t\left(\theta^* | \theta^{(t-1)}\right)$

3. Compute the acceptance ratio

$$\rho = \frac{f\left(\theta^*\right) / J_t\left(\theta^* \middle| \theta^{(t-1)}\right)}{f\left(\theta^{(t-1)}\right) / J_t\left(\theta^{(t-1)} \middle| \theta^*\right)} \tag{2.32}$$

4. Accept/reject:

$$\theta^t = \begin{cases} \theta^* & \text{with probability } \min(\rho, 1) \\ \theta^{(t-1)} & \text{otherwise} \end{cases}. \tag{2.33}$$

□

The Markov chain produced by the Metropolis–Hastings algorithm is guaranteed to have f as its stationary distribution.

Theorem 2.25. Let $\left\{\theta^{(t)}\right\}$ be the chain produced by the Metropolis–Hastings algorithm, and $\text{Supp}(f)$ the support of f, i.e. $\text{Supp}(f) = \{x | f(x) > 0\}$. Then for any proposal distribution whose support includes $\text{Supp}(f)$, the stationary distribution of the chain is f. □

However, as the convergence is only asymptotic, we are far more interested in algorithms that convergences efficiently. Unfortunately, the vanilla Metropolis–Hastings algorithm is known to exhibit random-walk behaviours, so in order to suppress such behaviours we resort to Hamiltonian Monte Carlo, which adds an auxiliary parameter called the "momentum". While efficient, as we will see in the definition of HMC, this method requires the gradient of the target, limiting the possible range of targets that can be simulated. Fortunately, for the vast majority of statistical models the gradient can be easily computed by applying the chain rule of differentiation to elementary functions.

Definition 2.26 (Hamiltonian Monte Carlo). The **Hamiltonian Monte Carlo (HMC)** is a hybrid Monte Carlo method, i.e. it combines MCMC with deterministic simulation methods. It proceeds as follows for all $t > 0$:

1. Draw the initial momentum parameter $r^{(0)}$ from a multivariate normal distribution $\mathcal{N}(\mathbf{0}, M)$, where M is the mass matrix.

2. Perform the "leapfrog" steps for L times:

   (a) Update the momentum parameter r using half-step of the gradient of the target

   $$r \leftarrow r + \frac{1}{2}\epsilon \nabla_\theta f(\theta). \tag{2.34}$$

   (b) Update the actual parameter θ

   $$\theta \leftarrow \theta + \epsilon M^{-1} r. \tag{2.35}$$

   (c) Use the gradient to half-update r

   $$r \leftarrow r + \frac{1}{2}\epsilon \nabla_\theta f(\theta). \tag{2.36}$$

3. Finally, apply the Metropolis–Hastings algorithm (Defintion 2.24) to the output of the L^{th} leapfrog step θ^* and r^*, with acceptance ratio

   $$\rho = \frac{\exp\left(f\left(\theta^*\right) - \frac{1}{2} r^* \cdot r^*\right)}{\exp\left(f\left(\theta^{(t-1)}\right) - \frac{1}{2} r^{(0)} \cdot r^{(0)}\right)}. \tag{2.37}$$

The number of leapfrog steps L, stepsize ϵ and the mass matrix M are parameters that need to be set for optimal performance, which can be done using an adaptive scheme such as NUTS (Hoffman et al., 2014). □

2.5 MCMC Convergence

After obtaining a sample of the target, we then need to assess if the MCMC outputs constitute a valid computation. We do this by checking if the Markov chain is indeed stationary. There are two key metrics of interest, the split-$\hat{R}$ and the effective sample size. From now on, we assume that the target f can be written as a parametric generative model with negative log-posterior $-\log[\mathbb{P}(\theta|y)]$ and that we have run M Markov chains each with N iterations, giving a total of $S = MN$ draws across all chains.

Definition 2.27 (Split-$\hat{R}$). Let $\theta^{(nm)}$ be the n^{th} draw of the m^{th} chain, $\overline{\theta}^{(\cdot m)}$ the average of draws from the m^{th} chain and $\overline{\theta}^{(\cdot\cdot)}$ the average for all draws. We now compute B and W, the **between-** and **within-chain variances**, as follows (Vehtari et al., 2021):

$$B = \frac{N}{M-1}\sum_{m=1}^{M}\left[\overline{\theta}^{(\cdot\cdot)} - \overline{\theta}^{(\cdot\cdot)}\right]^2, \text{where} \tag{2.38}$$

$$\overline{\theta}^{(\cdot m)} = \frac{1}{N}\sum_{n=1}^{N}\theta^{(mn)} \tag{2.39}$$

$$\overline{\theta}^{(\cdot\cdot)} = \frac{1}{N}\sum_{m=1}^{M}\overline{\theta}^{(\cdot m)}; \tag{2.40}$$

$$W = \frac{1}{M}\frac{1}{N-1}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\theta^{(nm)} - \overline{\theta}^{(\cdot m)}\right]^2. \tag{2.41}$$

The **scale reduction factor**, or $\hat{R}$, is then given by

$$\hat{R} = \sqrt{\frac{\hat{\mathbb{V}}(\theta|y)}{W}}, \text{where} \tag{2.42}$$

$$\hat{\mathbb{V}}(\theta|y) = \frac{N-1}{N}W + \frac{1}{N}B. \tag{2.43}$$

$\hat{R}$ represents the factor by which the scale of the current distribution of θ would reduced as $N \to \infty$. Finally, before computing $\hat{R}$ we split each of the M chain in half so as to minimize the influence of

non-stationary chains that cover a similar range of values (see Figure 1b in (Vehtari et al., 2021) for an example).

Another issue with MCMC samples is that they may be autocorrelated, which needs to be corrected. We use the concept of an effective sample size to formalize this.

Definition 2.28 (Effective Sample Size). The **effecitve sample size** (ESS) is a measure of how many *independent samples* there are for a MCMC sample. It is given by

$$S_{\text{eff}} = \frac{NM}{1 + 2 \sum_{t=1}^{T} \rho_t}, \tag{2.44}$$

where ρ_t is the autocorrelation for lag t. See (Vehtari et al., 2021) for how to compute ρ_t. □

For HMC sampler, it is usually possible to produce antithetic samples, that is, negatively correlated samples, so that the ESS may be larger than the number of MCMC samples. For the purpose of model diagnosis, (Vehtari et al., 2021) recommends $\hat{R} < 1.01$ for all variables of interest and $S_{\text{eff}} > 400$ based on 4 diffusely-intiated independent chains, or $S_{\text{eff}} > 100$ per chain if running more chains. By contrast, (Kruschke, 2015) recommends that $S_{\text{eff}} \geqslant 10000$ as an heuristic for stable estimation of tail quantities.

2.6 Model Evaluation

For model evaluation, we will use approximated leave-one-out (LOO) cross-validation via Pareto smoothed importance sampling (PSIS) (Vehtari et al., 2017), which computes the expected log pointwise predictive density (ELPD) for new data points, given by

$$\text{ELPD}_{\text{loo}} = \sum_{i=1}^{N} \log \underbrace{\int \mathbb{P}(y_i|\boldsymbol{\theta}) \mathbb{P}(\boldsymbol{\theta}|y_{-i}) \, d\boldsymbol{\theta}}_{\mathbb{P}(y_i|y_{i-1})}, \tag{2.45}$$

by drawing samples of $\boldsymbol{\theta}$ from the posterior using importance sampling with importance ratios

$$r_i^{(s)} = \frac{1}{\mathbb{P}\left(y_i|\boldsymbol{\theta}^{(s)}\right)}, \quad s \in [1, S] \cap \mathbb{Z}, \tag{2.46}$$

smoothing data points whose $r^{(s)}$ fall in the top 20% with the quantile function of the generalized Pareto distribution

$$\tilde{w}_i^{(s)} = F_{\text{Pareto}}^{-1}\left(\frac{z - 1/2}{0.2S}\right), \quad z \in [1, 0.2S] \cap \mathbb{Z} \tag{2.47}$$

truncating the weights $w_i^{(s)}$ at $S^{3/4}\tilde{w}_i$, where $\tilde{w}_i$ is the mean of the $\tilde{w}_i^{(s)}$'s over all $s \in [1, S] \cap \mathbb{Z}$, to ensure finite variance, and finally evaluating the integrand in (2.45)

$$\mathbb{P}(y_i|y_{i-1}) \approx \frac{S}{\sum_{s=1}^{S} r_i^{(s)}}. \tag{2.48}$$

Finally, the LOO ELPD estimator can be approximated with

$$\widehat{\text{ELPD}}_{\text{PSIS-LOO}} = \sum_{i=1}^{N} \log\left[\frac{\sum_{s=1}^{S} w_i^{(s)} \mathbb{P}\left(y_i|\,\theta^{(s)}\right)}{\sum_{s=1}^{S} w_i^s}\right], \tag{2.49}$$

where $w_i^{(s)}$ are the truncated weights above. However, if the shape parameter $\hat{k}$ of the fitted generalized Pareto distribution is greater than 0.7, the corresponding data point is considered "problematic" — it is an outlier that has a large influence on the ELPD — and if the number of problematic data points is large, K-fold cross-validation (CV) would be more robust than PSIS-LOO (Vehtari and Lampinen, 2002; Vehtari et al., 2017). To perform K-fold cross validation, the data is partitinoed

in to K subsets $\{y_k\}_{k=1}^{K}$ and compute the log predictive density for each fold given the other folds

$$\log \mathbb{P}\left(y_i \middle| y_{(-k)}\right) = \log \int \mathbb{P}\left(y_i \middle| \theta\right) \mathbb{P}\left(\theta \middle| y_{(-k)}\right) d\theta. \tag{2.50}$$

We will use $K = 10$ whenever LOO-CV is invalid and K-fold CV would have to be used.

CHAPTER 3

IBM MARKETSCAN

3.1 Cohort Construction

We used for our analysis the IBM Health MarketScan commercial insurance dataset (IBM Watson Health, 2019), covering health event of more than half of the US population during 2003–2018. The data provide person- and daily-level resolution of disease diagnoses, prescription medication, medical procedures, and family linkage information inferred from co-insurance data. Our analysis of these data enabled us to calculate mother-child links for over 3.1 million mother-child pairs.

Specifically, we matched newborns with their mothers using the following procedure. The newborn was identified as an enrollee 0 years and at least one live-delivery diagnostic code in the International Classification of Diseases (ICD, versions 9 and 10) (World Health Organization, 1992). The mother was identified as a female covered by the same family insurance policy as the newborn having at least one live-delivery-related diagnostic code registered within 3 days of the the day on which the first newborn delivery-related code was registered, which was treated in subsequent analyses as the birthday of the said newborn. For the sex ratio at birth study (Ch. 4), we used an earlier version of the data, but the procedure up to this point was the same. In that dataset, we found 3,133,062 newborns, of which 2,096,775 were matched to a mother. Our approach was notably different from that of (Messinger et al., 2020), where the objective was to track the menstrual periods of female enrollees irrespective of births.

For models concerning neurodevelopmental disorders (NDDs) (Ch. 5), since we are interested in the health profile of the newborns in early life, we required newborns to have been followed up immediately afterbirth with no more than 1 month of lapse between birth and first day of enrolment (n = 2,843,213), as in many cases the enrolment period started only on the first day of the month following the month of birth. As the minimum age at diagnosis of the NDDs is 1 (cf. (Straub et al., 2021)), we only have included only newborns with at least 1 year of postnatal follow-

up (n = 1,880,399). Moreover, since we are also interested in probing the effect of air pollution ($PM_{2.5}$), we have additionally subset individuals with geographical information encoded with Federal Information Processing Standard Publication (FIPS), which covered subjects with enrolment between 2003–2011 (n = 834,290). The process is summarized succinctly in Figure 3.1.

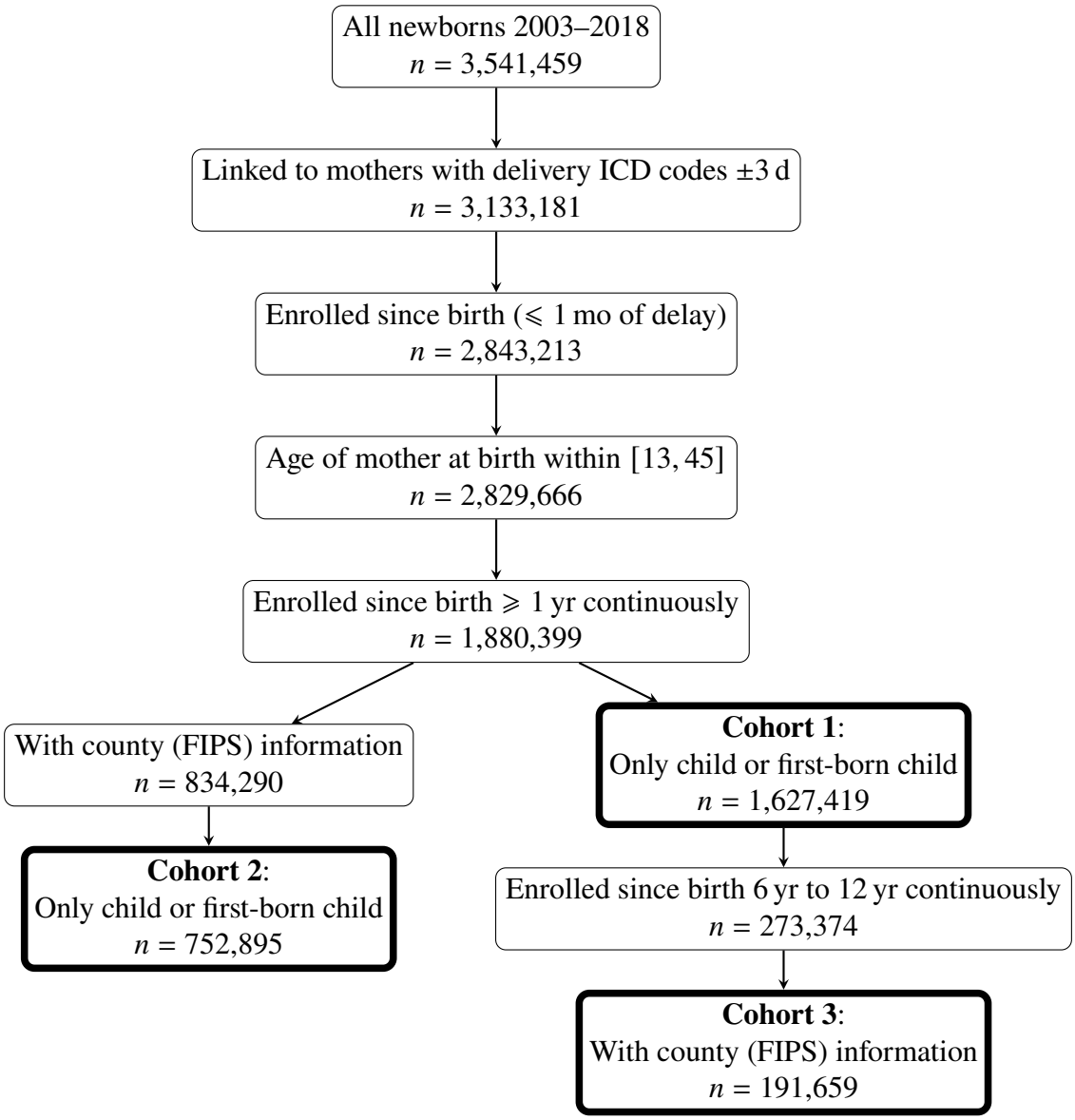


Figure 3.1: Flowchart for inclusion/exclusion of subjects. Nodes with increased line thickness indicated final cohorts used either in the main analysis or in some sensitivity analysis.

3.2 Summary Statistics

Tables 3.1 to 3.3 contain basic summary statistics of the 3 cohorts.

Risk Factor	Total	No MIA	MIA
Sex: Female	782755 (48.10%)	569047 (48.10%)	213708 (48.09%)
Premature Birth: 32 ~ 37 weeks	61411 (3.77%)	42518 (3.59%)	18893 (4.25%)
Very Premature Birth: < 32 weeks	25907 (1.59%)	18590 (1.57%)	7317 (1.65%)
Caesarean Section Delivery	575059 (35.34%)	405193 (34.25%)	169866 (38.23%)
Low Birth Weight: $\leqslant$ 2500 g	17958 (1.10%)	12350 (1.04%)	5608 (1.26%)
High Birth Weight $\geqslant$ 4500 g	4792 (0.29%)	3429 (0.29%)	1363 (0.31%)
Teenage Mother: 13 ~ 19 years old	7638 (0.47%)	4955 (0.42%)	2683 (0.60%)
Advanced Age Mother: 35 ~ 45 years old	446952 (27.46%)	328486 (27.77%)	118466 (26.66%)
Gestation weeks	277.54 ± 11.34 [160.00, 280.00]	277.62 ± 11.21 [160.00, 280.00]	277.33 ± 11.66 [160.00, 280.00]
Years of follow-up	3.76 ± 2.83 [1.00, 16.00]	3.80 ± 2.86 [1.00, 16.00]	3.66 ± 2.73 [1.00, 15.97]

Table 3.1: Summary Statistics of maternal risk factors for **Cohort 1**: (cf. Figure 3.1); MIA stands for maternal immune activation, defined as the presence of any diagnosis codes of at least one of the immune-system related disorders; counts are represented by raw counts and their percentage in parentheses "()"; continuous variables are represented by their mean ± standard deviation as well as range in square brackets "[]"

Risk Factor	Total	No MIA	MIA
Sex: Female	362347 (48.13%)	271215 (48.13%)	91132 (48.12%)
Premature Birth: 32 ~ 37 weeks	20401 (2.71%)	14771 (2.62%)	5630 (2.97%)
Very Premature Birth: < 32 weeks	11046 (1.47%)	8193 (1.45%)	2853 (1.51%)
Caesarean Section Devliery	264688 (35.16%)	192440 (34.15%)	72248 (38.15%)
Low Birth Weight: $\leqslant$ 2500 g	5401 (0.72%)	4006 (0.71%)	1395 (0.74%)
High Birth Weight $\geqslant$ 4500 g	2167 (0.29%)	1597 (0.28%)	570 (0.30%)
Teenage Mother: 13 ~ 19 years old	4506 (0.60%)	2963 (0.53%)	1543 (0.81%)
Advanced Age Mother: 35 ~ 45 years old	244372 (32.46%)	185376 (32.90%)	58996 (31.15%)
Gestation weeks	278.02 ± 10.39 [160.00, 280.00]	278.06 ± 10.35 [160.00, 280.00]	277.92 ± 10.52 [160.00, 280.00]
Years of follow-up	4.60 ± 3.39 [1.00, 16.00]	4.60 ± 3.40 [1.00, 16.00]	4.61 ± 3.35 [1.00, 15.96]
Mean $PM_{2.5}$	11.36 ± 2.34 [4.09, 21.29]	11.37 ± 2.35 [4.09, 21.29]	11.31 ± 2.30 [4.20, 20.36]
Maximum $PM_{2.5}$	12.29 ± 2.60 [4.33, 25.78]	12.31 ± 2.63 [4.33, 25.78]	12.23 ± 2.54 [4.33, 24.79]
Median $PM_{2.5}$	11.40 ± 2.35 [4.13, 22.05]	11.42 ± 2.37 [4.13, 22.05]	11.36 ± 2.31 [4.18, 20.55]
Population $PM_{2.5}$	11.35 ± 2.33 [4.08, 21.19]	11.36 ± 2.34 [4.08, 21.19]	11.31 ± 2.29 [4.25, 20.37]

Table 3.2: Summary Statistics of maternal risk factors for **Cohort 2** (cf. Figure 3.1); MIA stands for maternal immune activation, defined as the presence of any diagnosis codes of at least one of the immune-system related disorders; counts are represented by raw counts and their percentage in parentheses "()"; continuous variables are represented by their mean ± standard deviation as well as range in square brackets "[]"

Risk Factor	Total	No MIA	MIA
Sex: Female	78514 (48.11%)	58303 (48.05%)	20211 (48.30%)
Premature Birth: 32 ~ 37 weeks	3993 (2.45%)	2855 (2.35%)	1138 (2.72%)
Very Premature Birth: < 32 weeks	2330 (1.43%)	1755 (1.45%)	575 (1.37%)
Caesarean Section Delivery	58515 (35.86%)	42362 (34.91%)	16153 (38.60%)
Low Birth Weight: $\leqslant$ 2500 g	1139 (0.70%)	849 (0.70%)	290 (0.69%)
High Birth Weight $\geqslant$ 4500 g	398 (0.24%)	304 (0.25%)	94 (0.22%)
Teenage Mother: 13 ~ 19 years old	371 (0.23%)	252 (0.21%)	119 (0.28%)
Advanced Age Mother: 35 ~ 45 years old	69112 (42.35%)	52639 (43.38%)	16473 (39.36%)
Gestation weeks	278.13 ± 10.21 [160.00, 280.00]	278.14 ± 10.23 [160.00, 280.00]	278.09 ± 10.15 [160.00, 280.00]
Years of follow-up	8.59 ± 1.66 [6.00, 12.00]	8.59 ± 1.67 [6.00, 12.00]	8.59 ± 1.66 [6.00, 12.00]
Mean $PM_{2.5}$	11.40 ± 2.31 [4.18, 20.64]	11.40 ± 2.33 [4.18, 20.64]	11.40 ± 2.25 [4.38, 20.36]
Maximum $PM_{2.5}$	12.33 ± 2.56 [4.39, 25.40]	12.34 ± 2.59 [4.39, 25.40]	12.30 ± 2.49 [4.75, 24.79]
Median $PM_{2.5}$	11.45 ± 2.32 [4.18, 22.05]	11.45 ± 2.34 [4.18, 22.05]	11.45 ± 2.27 [4.36, 20.55]
Population $PM_{2.5}$	11.40 ± 2.30 [4.17, 20.73]	11.40 ± 2.32 [4.17, 20.73]	11.40 ± 2.25 [4.42, 20.37]

Table 3.3: Summary Statistics of maternal risk factors for **Cohort 3** (cf. Figure 3.1); MIA stands for maternal immune activation, defined as the presence of any diagnosis codes of at least one of the immune-system related disorders; counts are represented by raw counts and their percentage in parentheses "()"; continuous variables are represented by their mean ± standard deviation as well as range in square brackets "[]"

CHAPTER 4

OBSERVABLE VARIATIONS IN HUMAN SEX RATIO AT BIRTH

This chapter was adapted from (Long et al., 2021). See this link for author information and contributions.

4.1 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 $\frac{1}{2}$(Boklage, 2005). 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 $\frac{1}{2}$ (Orzack et al., 2015). 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 (Bruckner and Catalano, 2018). Third, SRB may be affected by peri-conceptual maternal hormonal levels (Grant and Chamley, 2010; James, 2013). 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 4.1).

While there are multiple studies which have observed the positive associations between air pol-

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 4.1: 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.

lution and spontaneous abortion (Leiser et al., 2019; Zhang et al., 2019a), most of those conclusions based on analyses of relatively small samples (Table 4.1), severely curtailing their statistical power. In this study, we harnessed the power of 2 very large datasets: the MarketScan insurance claim data (Hansen, 2017) 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 $\sim$ 3 million newborns from 1983 to 2013) (Emilsson et al., 2015). Our present study is the first systematic investigation of numerous chemical pollutants and other environmental factors using large datasets from two continents.

4.2 Methods

4.2.1 Data

The IBM Health MarketScan dataset (Hansen, 2017) represents 104,565,671 unique individuals and 3,134,062 unique live births. The Swedish National Patient Register (Emilsson et al., 2015) record health statistics for over ten million individuals, and 3,260,304 unique live births. We juxta-

posed 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 (EPA), the Swedish Meteorological and Hydrological Institute and 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 (Lobdell et al., 2011; Messer et al., 2014).

4.2.2 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 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.

4.2.3 Regression Analysis

We used multilevel Bayesian logistic regression with random effects implemented in the R-language package `rstan` (Stan Development Team, 2020). To facilitate model building, we used the R-language package `brms` (Bürkner, 2017) with default priors. Sampling was performed with the No-U-Turn sampler (NUTS) (Hoffman et al., 2014) with 500 warm-up steps and 1500 iteration steps with 28 Markov chains, of which the convergence was asseessed using the $\hat{R}$ statistic (Vehtari

et al., 2021). The model for the j^{th} factor (predictor) is given as follows:

$$\text{logit}\left(p_j\right) = \log\left(\frac{p_j}{1-p_j}\right) = \alpha_{[k]j} + \boldsymbol{\beta}_j^{\intercal}\mathbf{x}_j, \tag{4.1}$$

where p_j is the probability that a newborn is male, $\mathbf{x}_j$ is the vector representing the j^{th} factor, $\boldsymbol{\beta}_j$ their coefficients, and $\alpha_{[k]j}$ the intercept for the k^{th} 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

$$\alpha_j = \mu_a + \eta_j, \tag{4.2}$$

$$\eta_j \sim \mathcal{N}\left(0, \sigma_\eta^2\right) \tag{4.3}$$

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

$$\alpha_j = \mu_a + \eta_j + \nu_j, \tag{4.4}$$

$$\eta_j \sim \mathcal{N}\left(0, \sigma_\eta^2\right), \tag{4.5}$$

$$\nu_j \sim \mathcal{N}\left(0, \sigma_\nu^2\right), \tag{4.6}$$

where η_j and ν_j are independent of each other and for all j (Gelman and Hill, 2006). 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 (Khan et al., 2019).

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 infor-

mation criterion relative to the null model (where $\mathbf{x}_j = \mathbf{0}$, i.e. 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 (Vehtari et al., 2017), and used the Benjamini–Yekutieli method to adjust for multiple comparisons (Benjamini et al., 2001).

4.2.4 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 (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. An sARIMA model is given by

$$\begin{aligned} &\underbrace{\left(1 - \sum_{j=1}^{P} \Phi_j L^j\right)}_{\text{sAR}} \underbrace{\left[\left(1 - L^S\right)^D\right]}_{\text{sI}} \underbrace{\left(1 - \sum_{j=1}^{p} \phi_j L^j\right)}_{\text{AR}} \underbrace{\left[(1 - L)^d\right]}_{\text{I}} y_t \\ &= \underbrace{\left(1 + \sum_{j'=1}^{Q} \Theta_{j'} L^{j'}\right)}_{\text{sMA}} \underbrace{\left(1 + \sum_{j'=1}^{q} \theta_{j'} L^{j'}\right)}_{\text{MA}} \varepsilon_t, \end{aligned} \tag{4.7}$$

where AR indicates the autoregression 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` (Hyndman and Khandakar, 2008; Hyndman et al., 2019) 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) (Schwarz,

1978). 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 (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):

$$y_t = Z_t^\intercal \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, H_t), \tag{4.8}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \qquad \eta_t \sim \mathcal{N}(0, Q_t), \tag{4.9}$$

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):

$$y_t = \mu_t + \tau_t + \varepsilon_t \tag{4.10}$$

$$\mu_t = \mu_t + \delta_{t-1} + u_t \tag{4.11}$$

$$\delta_t = \delta_{t-1} + v_t \tag{4.12}$$

$$\tau_t = -\sum_{s=1}^{S-1} \tau_{t-s} + w_t. \tag{4.13}$$

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 4.10–4.13 hold. We used the R package `CausalImpact` (Brodersen et al., 2015), which in turn relied on the R package `bsts` (Scott, 2019) as backend, to fit the data.

4.2.5 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).

4.3 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 (Lerchl, 1998; Melnikov and Grech, 2003). 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 (Catalano et al., 2008; Helle et al., 2009; Fukuda et al., 2014; Dixson et al., 2011), 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 4.1**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 4.2 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.

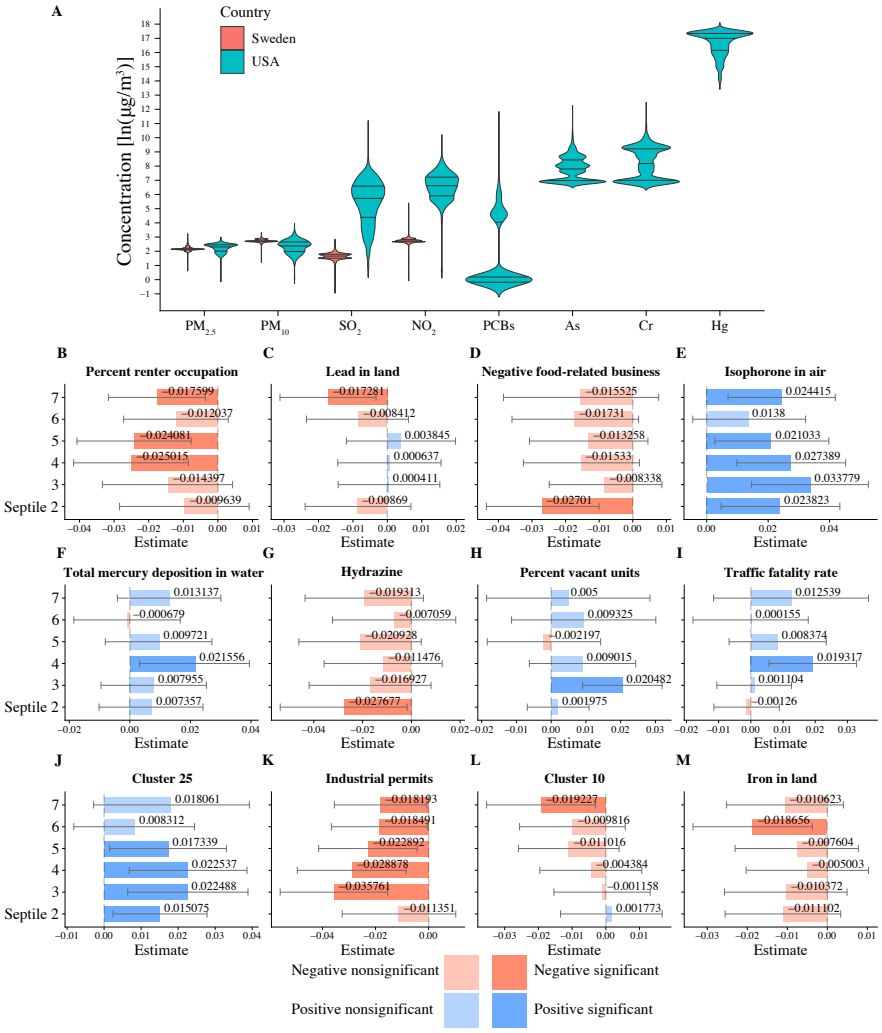


Figure 4.1: Airborne health-related substances and their association with the SRB. **A**: Comparison of airborne pollutant concentrations across the US (cyan violin plots) and Sweden (pink violin plots). Only 4 air components, fine particulate matter ($PM_{2.5}$), coarse particulate matter (PM_{10}), 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 4.2. 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 4.2: 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 4.3). Specifically, our data suggests that 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 4.3. We also established several new environmental associations that have not been reported before (Table 4.4, Fig 4.1 **B–M**, and Fig 4.2). Fig 4.1 **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). In the case of PCBs (polychlorinated biphenyls), on which the literature has reported conflicting evidences (Pavic, 2020), we found a positive correlation with the SRB. Since the sample sizes of the studies published thus far were very small (cf. Table 1), our PCBs result would have substantially larger statistical power.

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 4.3: Test results for factors selected from the literature reports (Table 4.1). We included a factor only if both its ΔIC and the coefficient of at least one of its septiles was statistically significant.

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 4.4: 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.

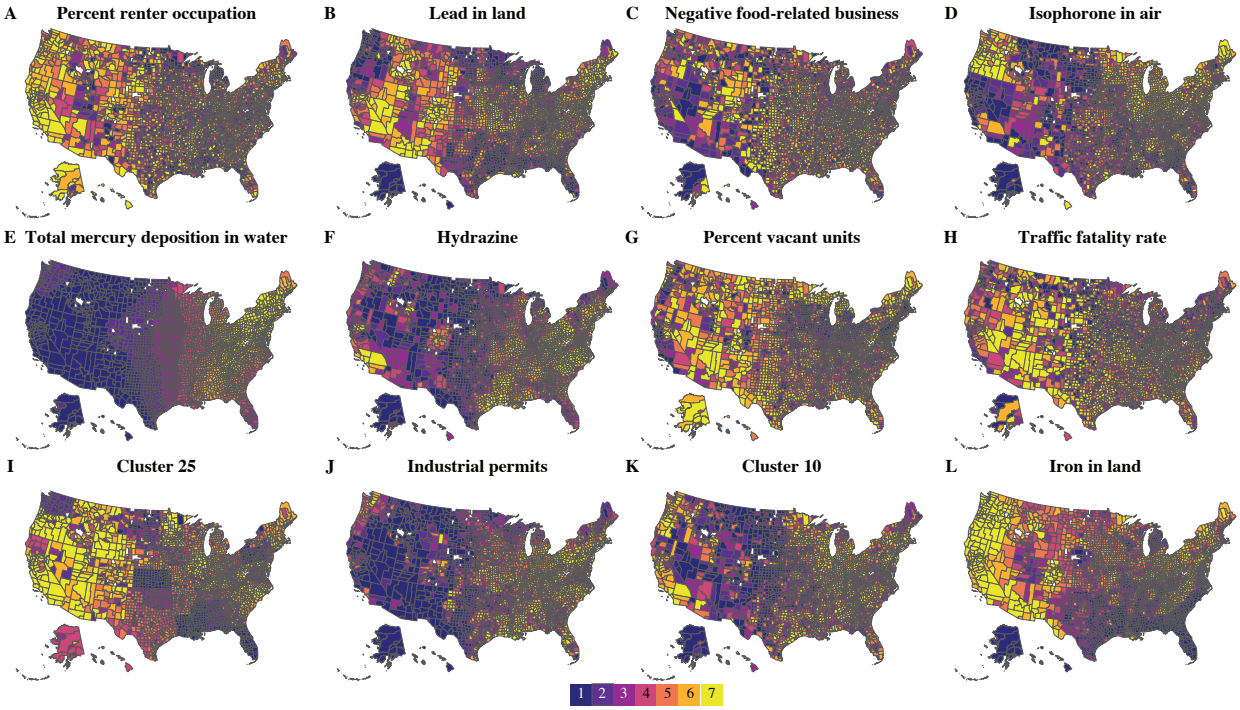


Figure 4.2: County-level geographical septile distribution for the first 12 statistically significant factors with at least one statistically significant coefficient ranked by decreasing ΔIC. The factors labelled **A–M** are the same as shown in Fig 4.1, Plates **B–M** and are ordered identically in both figures.

The geographic distribution of these pollutants varies remarkably, as seen in Fig 4.2. For example, lead in land (Fig 4.2**B**) appears to be enriched in the northeast, southwest, and mid-east US, but not in the south. Hydrazine (Fig 4.2**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 4.2**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 4.3C and 4.4C, cf. Tables S5c and S6c).

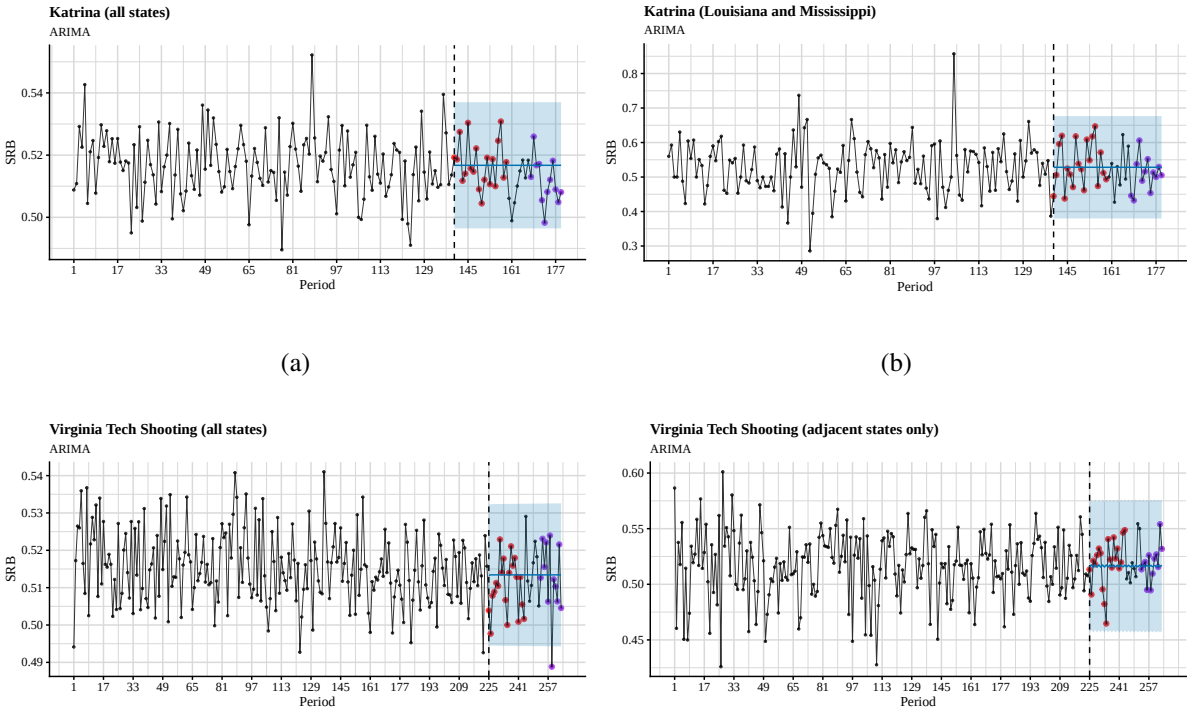


Figure 4.3: Time series plots and out-of-sample forecasts for SRB data grouped into 7-day periods and fitted with seasonal ARIMA models. The blue shade is the 95% confidence level. The observed SRBs for the first five months after the intervention are presented by red dots, whereas the observed SRBs 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.

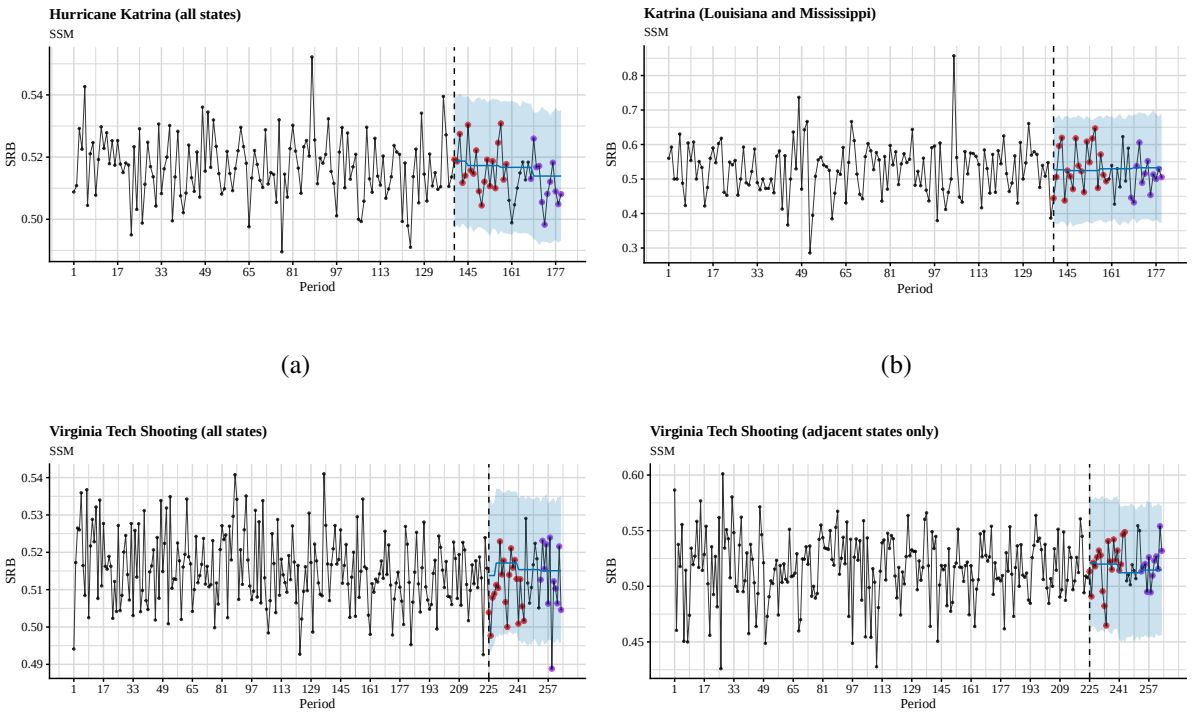


Figure 4.4: Time series plots and out-of-sample forecasts for SRB data grouped into 7-day periods and fitted with state space models. The blue shade is the 95% confidence level. The observed SRBs for the first five months after the intervention are presented by red dots, whereas the observed SRBs for 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.

4.4 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 (Douhard, 2017). The most frequently mentioned theory is the Trivers–Willard hypothesis (TWH), named after the researchers who proposed it (Trivers and Willard, 1973)). The TWH postulates that, because the cost of rearing children is much higher for females than for males, in favourable, resource-rich environments, males would 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 (Trivers and Willard, 1973; James, 2006, 2013):

Proposition 4.1. 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.

Proposition 4.2. A2 The condition of the offspring persists after parental investment ends, and is positively correlated with the offspring's reproductive success.

Proposition 4.3. 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 4.4. C1 The SRB varies such that females in favourable conditions have more male offspring, and in unfavourable conditions, more female offspring.

Assumption 4.3 is called Bateman's principle (Bateman, 1948) (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 (Hoquet et al., 2020; Gowaty et al., 2012) for critiques of Bateman's methodology). Nevertheless, a modified version of BP, which generalizes anisogamy to parental investment, has enjoyed prominence among evolutionary biologists (Trivers, 1972). 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 (Hrdy, 1986). 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 (Hubbard, 1990; Gerlach et al., 2012).

Faced with such criticisms as well as an increasing amount of evidence from species across the animal kingdom that did not conform to BP (Tang-Martínez, 2016), supporters of BP have responded that sex differences ultimately originated from *historical* anisogamy (Schärer et al., 2012; Kokko et al., 2013), 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 (Morimoto, 2020; De Lisle, 2019). Moreover, as a counter-challenge to the former point, supporters of BP 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 (Janicke et al., 2016). This reworking allowed for a potential remedy for BP, namely by generalizing it as follows (Arnold and Duvall, 1994; Tang-Martínez, 2016; Hoquet, 2020b):

Proposition 4.5. 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 4.6. C1* The SRB should vary such that females in favourable conditions have more offspring of the sex more susceptible to sexual selection, and in unfavourable conditions, more offspring of the 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 (Morimoto, 2020). Candidates for the identity of that sex include higher variance in number of (adult) offspring and higher variance in parental investments (James, 2008, 2013). Nevertheless, under this revised framework, for sexually dimorphic selection patterns to develop and persist as opposed to randomly fluctuating across time (Gowaty and Hubbell, 2009), one inevitably has to invoke the sexual cascade hypothesis: a small initial difference (e.g. anisogamy) in sex-related phenotype will "snowball" into larger, persistent patterns through hereditary feedback loops (Parker, 2014; Parker and Pizzari, 2015; Fromhage and Jennions, 2016). Such cascading has also featured in the above-mentioned meta-analysis discussion regarding high-level explanatory patterns among animal species (Janicke et al., 2016), bracketing all differences in sex-related traits into one-dimensional sexual selection (Tang-Martínez, 2016). There is a plethora of other competing theories, e.g. (Gowaty and Hubbell, 2009) and (Hoquet, 2020b), which predict largely stochastic variations of sex-related phenotypes, emphasizing the role social and ecological factors have played in shaping *plastic* sex-roles (Tang-Martínez, 2016; Gowaty, 2015; Roughgarden, 2015). Even if the last point may still be somewhat contentious (Parker and Pizzari, 2015), 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 (Fine, 2017, 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 (James, 2006), 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 (Zietsch et al., 2020; Catalano et al., 2020). In particular, Zietsch et al. (2020) have demonstrated that there exists neither within-individual SRB auto-correlation (contra Assumption 4.1) nor similarity in the SRB for children of siblings (contra Assumption 4.2). 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. 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 (James, 2008), maternal dominance hypothesis (Grant, 2003, 2007) and the Bruce effect (Catalano et al., 2018) in the same predicament. Appealing to evolutionary history (i.e. TWH was in operation in the past but not at present, or TWH is an effect of some vestigial evolutionary mechanism) is of no help here, since an adaptive selection mechanism cannot explain why and how, at some point in history, the heritability was lost (Zietsch et al., 2021). In other words, if SRB is ever influenced by some factor(s) at least partially heritable, then SRB itself would have to be heritable as well, which the results from Zietsch et al. rule out. Thus, our results are better interpreted as supporting the overwhelming influence of random Mendelian segregation on the SRB (cf. (Postma et al., 2011) which claims *complete* attribution of SRB variation to Mendelian segregation in some non-human species), such that SRB variations are at least primarily due to *non-adaptive* (e.g. socio-cultural (Zietsch et al., 2020; Dupré, 2012, Ch. 14)) 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 (Bruckner and Catalano, 2018), 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 James and Grech (2018)), *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 (Terrell et al., 2011; James and Grech, 2017), and that our results do support the non-monotonous, dose-response profiles frequently reported in the literature (Pavic, 2020) (Table 4.3). 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 (Catalano et al., 2020) using large datasets with covariates sampled frequently across considerable spatio-temporal ranges (Zietsch et al., 2020). Another interesting direction would be to determine the potentially non-adaptive physiological mechanisms.

4.5 Limitations

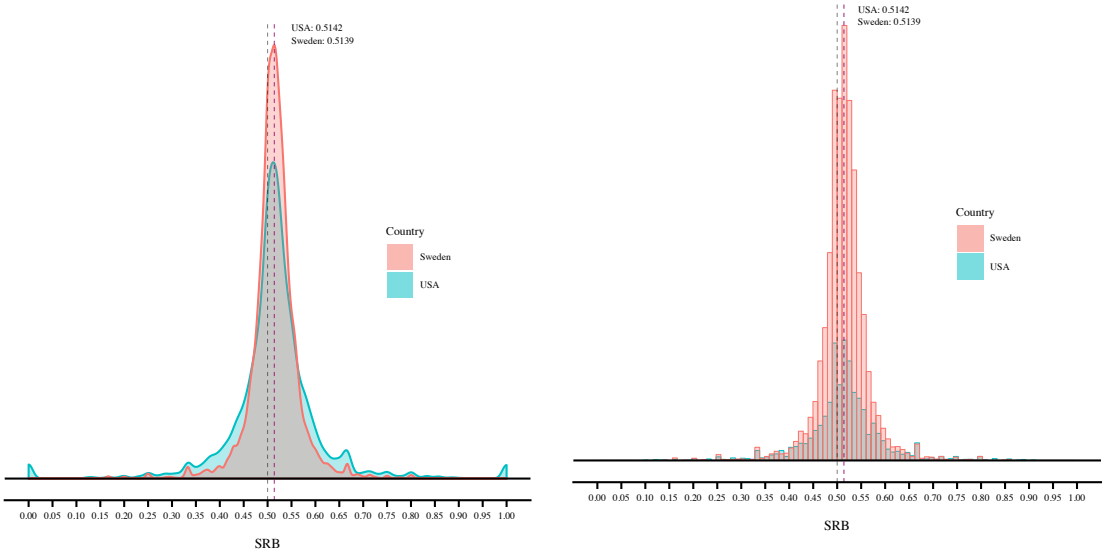
Unlike some of the recent studies (Catalano et al., 2005), we did not have access to the sex of stillbirths, which would have enabled us to probe negative selection *in utero* against frail males (Bruckner and Catalano, 2018). When quantifying pollutants in the US, we used the EPA air quality raw data, 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.

4.6 Appendix

4.6.1 Overall SRB Distribution

Figure 4.5 shows that the distributions of sex ratio at the county level (USA) or the kommun level (Sweden) are very similar, with the US having an overall SRB of 0.5142 and Sweden 0.5139.



(a) Kernel density estimation (KDE)

(b) Histogram

Figure 4.5: Distribution of the SRB in the US and Sweden at the county level (US) or the kommun level (Sweden)

4.6.2 Cluster Analysis

Figure 4.6 shows the dendrogram of the clustering the factors in the US EQI data set by Ward's method (see the Methods section in the main text for more detail). Each red box delimits a statistically significant cluster (at the 95% level), which contains at least 2 factors.

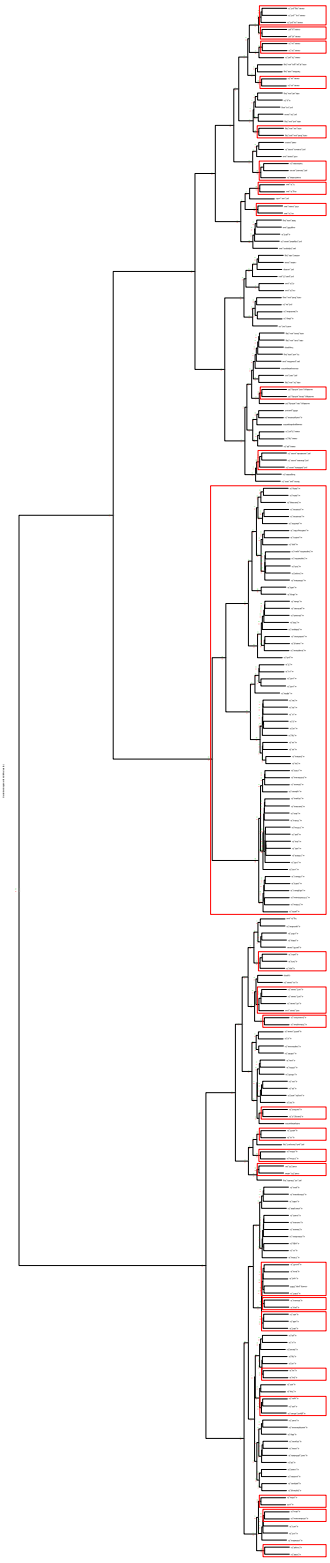


Figure 4.6: Dendrogram with statistically significant clusters (95% level) in red boxes

4.6.3 Regression

Tables 4.5 and 4.6, respectively, list all the statistically significant factors (8 for fixed-effect and all 125 for mixed-effect), sorted by decreasing ΔIC, with respect to the null.

Factor	ΔIC	SE
cluster_ward_8	25.6880	7.7901
cluster_ward_1	21.5759	6.4210
cluster_ward_11	20.9930	5.1878
cluster_ward_15	20.8578	4.1703
sewagenpdesperkm	19.0802	4.0902
pct_no_eng	16.1841	4.2216
a_dbp_ln	13.9700	3.8312
pct_mt_10units_log	10.5599	3.3704

Table 4.5: Differences in information criterion (ΔIC) and their standard errors (SE) of individual factors with fixed-effect only. Non-significant factors are omitted.

Table 4.6: Differences in information criterion (ΔIC) and their standard errors (SE) of individual factors with the random effect at the state level in the US EQI dataset.

Factor	ΔIC	SE
pct_rent_occ	54.2185	8.7760
cluster_ward_9	52.3907	8.8256
mean_pb_ln	51.7339	9.6749
farms_per_acre_ln	51.6266	9.0878
cluster_ward_15	51.5511	8.5623
pct_mt_10units_log	51.3259	8.5483
rate_food_env_neg	51.3230	9.2907
a_isophorone_ln	50.9430	10.3995
rate_ent_env_log	50.9259	8.8088

Continued on next page

Factor	ΔIC	SE
a_dbp_ln	50.4537	8.6102
a_mn_ln	50.2819	8.1831
hg_ln_ave	50.0576	8.6584
ryprop	49.6911	9.4066
a_n2h2_ln	49.5381	8.4877
pct_vac_units	49.3265	8.5257
fatal_rate_log	49.0348	8.3392
cluster_ward_25	48.6212	9.8838
indnpdesperkm	48.5428	10.2509
a_biphenyl_ln	48.4603	8.6128
a_cn_ln	48.1418	9.7888
cluster_ward_10	47.7198	8.7054
work_out_co	47.4916	8.9555
mean_fe_pct_ln	47.2117	8.0750
cluster_ward_19	47.1722	8.5983
cluster_ward_26	47.1691	8.7901
a_quinoline_ln	47.1477	9.6906
a_sb_ln	47.0255	8.3046
cluster_ward_8	46.3741	8.9328
med_rooms	46.3220	8.1562
rate_al_pn_gm_env_log	46.3080	8.3765
no3_mean_ave	46.1843	8.8100

Continued on next page

Factor	ΔIC	SE
a_meoh_ln	45.9854	8.3145
cluster_ward_6	45.9191	8.6963
d303_percent	45.8835	9.4942
na_ln_ave	45.0601	8.1385
a_co_mean_ln	44.9827	8.5933
a_c6h5cl_ln	44.6342	8.7028
a_me2_phthalate_ln	44.4614	8.0751
a_acrylic_acid_ln	44.2292	9.8383
a_mecl_ln	44.0207	8.4892
cat	43.9278	8.3540
mean_ti_pct_ln	43.8613	8.8678
cluster_ward_18	43.8523	9.3605
a_c2hcl3_ln	43.7136	8.7911
a_pahpom_ln	43.6743	9.1561
a_benzidine_ln	43.6200	8.4935
rate_civic_env_log	43.6027	8.5611
fungicides_ln	43.5278	8.7614
cluster_ward_3	43.5188	8.9542
pct_irrigated_acres_ln	43.4427	9.8728
a_p_ln	43.2101	8.2484
to_unit_rate_log	42.9314	8.5853
cluster_ward_2	42.8250	9.4466

Continued on next page

Factor	ΔIC	SE
pct_au_ln	42.8249	9.4111
cluster_ward_11	42.8243	9.0778
a_c3h3n_ln	42.8093	8.4624
a_diesel_ln	42.7350	8.1946
a_etacrylate_ln	42.7323	8.7796
avgofd3_ave	42.5189	8.8904
pct_pub_transport_log	42.1074	8.4142
a_pb_ln	41.9983	8.0841
a_ph3_ln	41.9038	9.0259
a_eox_ln	41.8434	8.9989
pct_unemp	41.8191	7.7675
a_hexane_ln	41.8044	8.0627
pct_no_eng	41.7541	8.2477
a_acrolein_ln	41.6847	8.2058
a_ech_ln	41.6194	8.4322
a_cumene_ln	41.6100	8.2629
a_11dce_ln	41.5146	8.4598
radon_zone	41.4107	8.6552
a_2np_ln	41.2805	9.3701
a_cr_ln	41.1006	8.3650
cluster_ward_13	41.0139	8.1444
a_pm10_mean_ln	40.8277	8.0920

Continued on next page

Factor	ΔIC	SE
a_cresol_ln	40.7598	8.2679
rate_trans_env_log	40.7569	8.4773
a_proo_ln	40.7404	8.1918
per_totpopss_ave	40.6453	8.2117
a_cl_ln	40.6386	8.2518
stormnpdesperkm	40.5914	8.4561
facilities_rate_log	40.5648	9.2041
pct_manure_acres_ln	40.4699	8.8810
cluster_ward_12	40.4274	9.1367
a_mibk_ln	40.3786	8.7446
a_dehp_ln	40.3305	8.1627
numdays_rain_activity_tot	40.1392	8.3155
cluster_ward_20	39.8644	8.6439
cl_ln_ave	39.5243	9.2021
a_egly_ln	39.4799	8.2135
mean_hg_ln	39.4692	8.1523
rate_ed_env_log	39.3155	8.7520
a_pm25_mean	39.1712	8.5654
cluster_ward_22	39.1668	8.3645
a_hg_ln	39.0513	7.8800
cluster_ward_5	39.0485	8.1282
a_chloroform_ln	38.9772	8.1250

Continued on next page

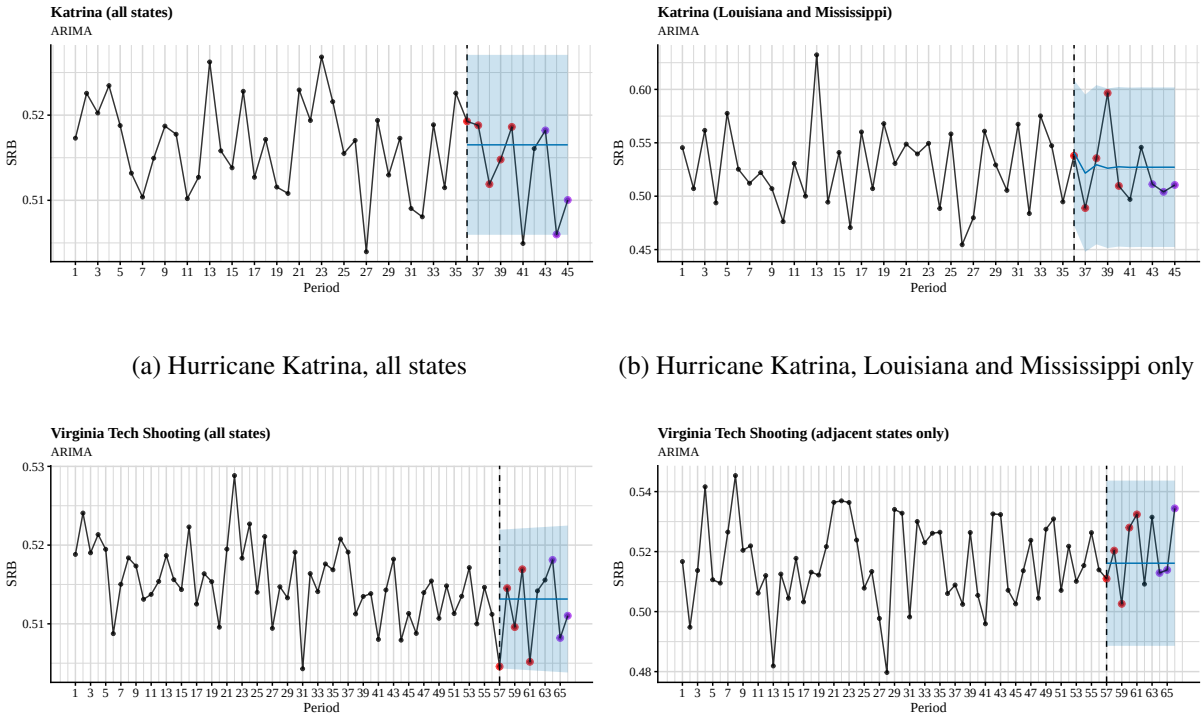
Factor	ΔIC	SE
a_cs2_ln	38.4509	7.6905
a_tdi_ln	38.4377	8.9270
per_pswithsw_ave	38.3795	8.1011
a_acetophenone_ln	37.9548	8.6551
violent_rate_log	37.8928	8.8251
a_cd_ln	37.8764	7.9860
a_dbcp_ln	37.7606	8.3128
a_etcl_ln	37.6943	8.3588
a_chloroprene_ln	37.6642	8.4969
cluster_ward_14	37.4994	8.6078
hwyprop	37.3042	8.4589
a_stryene_ln	37.2645	8.3137
cluster_ward_4	37.0974	8.2277
cluster_ward_1	36.9302	8.3875
cluster_ward_16	36.9113	8.0581
cluster_ward_7	36.6871	7.8672
rate_hc_env_log	36.6085	8.1125
cluster_ward_23	36.5280	8.1485
sewagenpdesperkm	36.0821	7.4545
cluster_ward_21	35.9375	8.7625
a_otoluidine_ln	35.8421	7.9246
a_mma_ln	35.3633	8.1445

Continued on next page

Factor	ΔIC	SE
a_mehydrazine_ln	35.1223	7.4720
cluster_ward_24	35.0831	8.0009
pct_pers_lt_pov	34.9108	7.1347
pct_hs_more	34.8201	8.0671
cluster_ward_17	34.5813	7.9217
a_hcl_ln	34.1790	8.0501

4.6.4 Time Series Forecasts

Monthly



(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

Figure 4.7: Time series plots and out-of-sample forecasts for SRB data grouped into 28-day periods and fitted with seasonal ARIMA models. The blue shade is the 95% confidence level. The observed SRBs for the first 5 months after the intervention are presented by red dots, whereas the observed SRBs for 7–9 months after the intervention are presented by purple dots. See also Table 4.7.

Period	SRB	Lower 95%	Upper 95%
36	0.5193	0.5060	0.5271
37	0.5188	0.5060	0.5271
38	0.5119	0.5060	0.5271
39	0.5148	0.5060	0.5271
40	0.5186	0.5060	0.5271
41(*)	0.5049	0.5060	0.5271
42	0.5161	0.5060	0.5271
43	0.5182	0.5060	0.5271
44	0.5060	0.5060	0.5271
45	0.5100	0.5060	0.5271

(a) Hurricane Katrina in all states

Period	SRB	Lower 95%	Upper 95%
36	0.5379	0.4726	0.6086
37	0.4889	0.4479	0.5952
38	0.5356	0.4549	0.6040
39	0.5966	0.4515	0.6009
40	0.5096	0.4528	0.6023
41	0.4970	0.4522	0.6017
42	0.5458	0.4525	0.6019
43	0.5112	0.4524	0.6018
44	0.5041	0.4524	0.6019
45	0.5105	0.4524	0.6019

(b) Hurricane Katrina in Louisiana and Mississippi

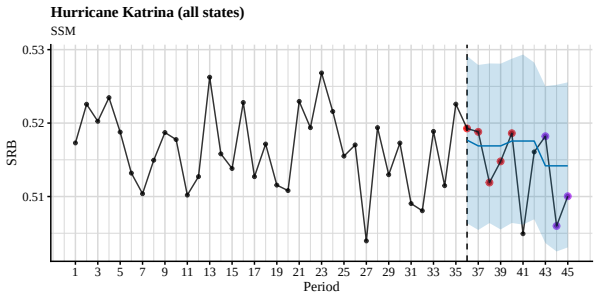
Period	SRB	Lower 95%	Upper 95%
57	0.5046	0.5043	0.5220
58	0.5145	0.5043	0.5220
59	0.5096	0.5042	0.5221
60	0.5169	0.5042	0.5221
61	0.5052	0.5041	0.5222
62	0.5142	0.5040	0.5222
63	0.5156	0.5040	0.5223
64	0.5181	0.5039	0.5224
65	0.5082	0.5039	0.5224
66	0.5110	0.5038	0.5225

(c) Virginia Tech Shooting in all states

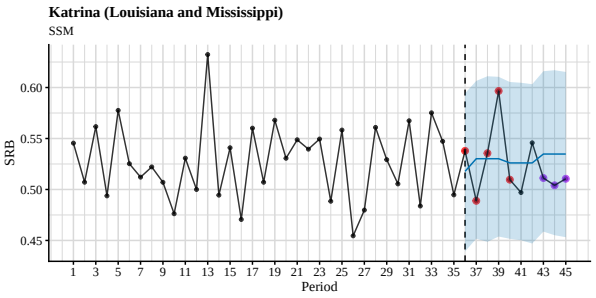
Period	SRB	Lower 95%	Upper 95%
57	0.5110	0.4886	0.5437
58	0.5203	0.4886	0.5437
59	0.5026	0.4886	0.5437
60	0.5280	0.4886	0.5437
61	0.5324	0.4886	0.5437
62	0.5092	0.4886	0.5437
63	0.5315	0.4886	0.5437
64	0.5129	0.4886	0.5437
65	0.5139	0.4886	0.5437
66	0.5344	0.4886	0.5437

(d) Virginia Tech Shooting in adjacent states

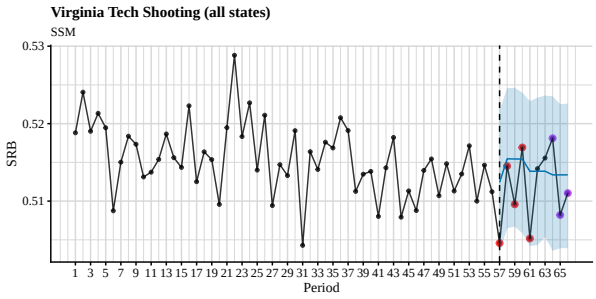
Table 4.7: Out-of-sample forecasts for the first 10 months after the intervention using SRB data grouped into 28-day periods and fitted with seasonal ARIMA models. Any period of which the observed SRB is outside of the 95% confidence level is marked by an asterisk (*).



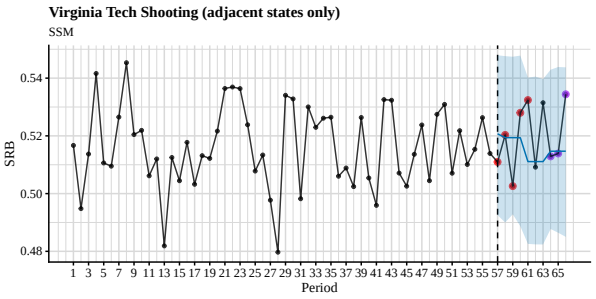
(a) Hurricane Katrina, all states



(b) Hurricane Katrina, Louisiana and Mississippi only



(c) Virginia Tech shooting, all states



(d) Virginia Tech shooting, adjacent states only

Figure 4.8: Time series plots and out-of-sample forecasts for SRB data grouped into 28-day periods and fitted with state-space models. The blue shade is the 95% confidence level. The observed SRBs for the first 5 months after the intervention are presented by red dots, whereas the observed SRBs for 7–9 months after the intervention are presented by purple dots. See also Table 4.8.

Period	SRB	Lower 95%	Upper 95%
36	0.5193	0.5063	0.5291
37	0.5188	0.5054	0.5279
38	0.5119	0.5064	0.5281
39	0.5148	0.5055	0.5281
40	0.5186	0.5064	0.5288
41	0.5049	0.5062	0.5293
42	0.5161	0.5069	0.5282
43	0.5182	0.5037	0.5250
44	0.5060	0.5025	0.5252
45	0.5100	0.5031	0.5255

(a) Hurricane Katrina in all states

Period	SRB	Lower 95%	Upper 95%
36	0.5379	0.4392	0.5944
37	0.4889	0.4517	0.6065
38	0.5356	0.4486	0.6111
39	0.5966	0.4538	0.6105
40	0.5096	0.4516	0.6054
41	0.4970	0.4502	0.6047
42	0.5458	0.4470	0.6033
43	0.5112	0.4587	0.6160
44	0.5041	0.4551	0.6170
45	0.5105	0.4532	0.6153

(b) Hurricane Katrina in Louisiana and Mississippi

Period	SRB	Lower 95%	Upper 95%
57	0.5046	0.5034	0.5217
58	0.5145	0.5065	0.5246
59	0.5096	0.5067	0.5246
60	0.5169	0.5059	0.5240
61	0.5052	0.5042	0.5229
62	0.5142	0.5043	0.5234
63	0.5156	0.5054	0.5236
64	0.5181	0.5036	0.5235
65	0.5082	0.5039	0.5225
66	0.5110	0.5039	0.5226

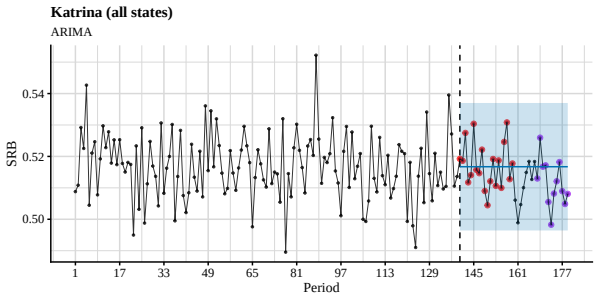
(c) Virginia Tech Shooting in all states

Period	SRB	Lower 95%	Upper 95%
57	0.5110	0.4924	0.5487
58	0.5203	0.4910	0.5478
59	0.5026	0.4926	0.5469
60	0.5280	0.4896	0.5481
61	0.5324	0.4832	0.5382
62	0.5092	0.4815	0.5390
63	0.5315	0.4829	0.5399
64	0.5129	0.4894	0.5432
65	0.5139	0.4866	0.5450
66	0.5344	0.4846	0.5433

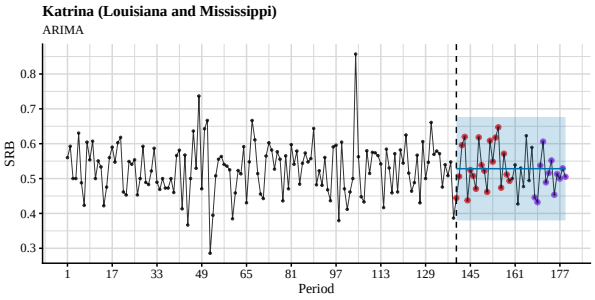
(d) Virginia Tech Shooting in adjacent states

Table 4.8: Out-of-sample forecasts for the first 10 months after the intervention using SRB data grouped into 28-day periods and fitted with state-space models. Any period of which the observed SRB is outside of the 95% confidence level is marked by an asterisk (*).

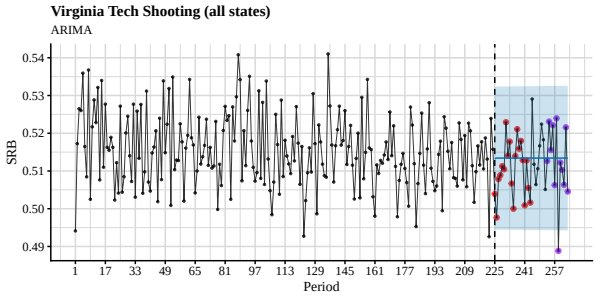
Weekly



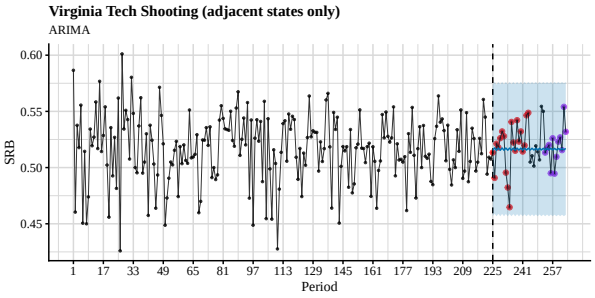
(a) Hurricane Katrina, all states



(b) Hurricane Katrina, Louisiana and Mississippi only



(c) Virginia Tech shooting, all states



(d) Virginia Tech shooting, adjacent states only

Figure 4.9: 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 5 months after the intervention are presented by red dots, whereas the observed SRBs for 7–9 months after the intervention are presented by purple dots. See also Table 4.9.

Period	SRB	Lower 95%	Upper 95%
140	0.5192	0.4964	0.5370
141	0.5186	0.4964	0.5370
142	0.5275	0.4964	0.5370
143	0.5117	0.4964	0.5370
144	0.5140	0.4964	0.5370
145	0.5304	0.4964	0.5370
146	0.5157	0.4964	0.5370
147	0.5147	0.4964	0.5370
148	0.5222	0.4964	0.5370
149	0.5090	0.4964	0.5370
150	0.5045	0.4964	0.5370
151	0.5121	0.4964	0.5370
152	0.5192	0.4964	0.5370
153	0.5106	0.4964	0.5370
154	0.5187	0.4964	0.5370
155	0.5100	0.4964	0.5370
156	0.5246	0.4964	0.5370
157	0.5308	0.4964	0.5370
158	0.5128	0.4964	0.5370
159	0.5177	0.4964	0.5370
160	0.5061	0.4964	0.5370
161	0.4989	0.4964	0.5370
162	0.5047	0.4964	0.5370
163	0.5101	0.4964	0.5370
164	0.5149	0.4964	0.5370
165	0.5184	0.4964	0.5370
166	0.5127	0.4964	0.5370
167	0.5184	0.4964	0.5370
168	0.5130	0.4964	0.5370
169	0.5260	0.4964	0.5370
170	0.5167	0.4964	0.5370
171	0.5171	0.4964	0.5370
172	0.5055	0.4964	0.5370
173	0.4983	0.4964	0.5370
174	0.5082	0.4964	0.5370
175	0.5121	0.4964	0.5370
176	0.5182	0.4964	0.5370
177	0.5089	0.4964	0.5370
178	0.5049	0.4964	0.5370
179	0.5081	0.4964	0.5370

(a) Hurricane Katrina in all states

Period	SRB	Lower 95%	Upper 95%
140	0.4444	0.3799	0.6767
141	0.5060	0.3799	0.6767
142	0.5957	0.3799	0.6767
143	0.6197	0.3799	0.6767
144	0.4375	0.3799	0.6767
145	0.5246	0.3799	0.6767
146	0.5077	0.3799	0.6767
147	0.4706	0.3799	0.6767
148	0.6182	0.3799	0.6767
149	0.5385	0.3799	0.6767
150	0.5224	0.3799	0.6767
151	0.4615	0.3799	0.6767
152	0.6087	0.3799	0.6767
153	0.5484	0.3799	0.6767
154	0.6176	0.3799	0.6767
155	0.6471	0.3799	0.6767
156	0.4737	0.3799	0.6767
157	0.5714	0.3799	0.6767
158	0.5116	0.3799	0.6767
159	0.4933	0.3799	0.6767
160	0.5000	0.3799	0.6767
161	0.5393	0.3799	0.6767
162	0.4271	0.3799	0.6767
163	0.5301	0.3799	0.6767
164	0.4769	0.3799	0.6767
165	0.6230	0.3799	0.6767
166	0.4941	0.3799	0.6767
167	0.5895	0.3799	0.6767
168	0.4458	0.3799	0.6767
169	0.4321	0.3799	0.6767
170	0.5376	0.3799	0.6767
171	0.6061	0.3799	0.6767
172	0.4891	0.3799	0.6767
173	0.5158	0.3799	0.6767
174	0.5521	0.3799	0.6767
175	0.4535	0.3799	0.6767
176	0.5132	0.3799	0.6767
177	0.5000	0.3799	0.6767
178	0.5294	0.3799	0.6767
179	0.5052	0.3799	0.6767

(b) Hurricane Katrina in Louisiana and Mississippi

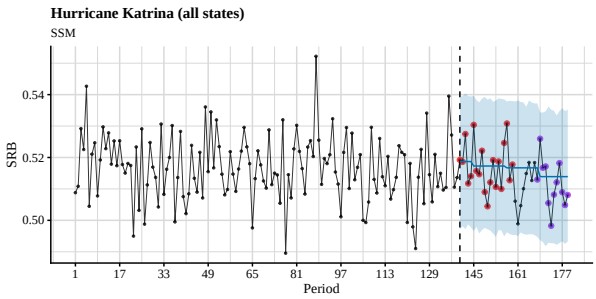
Period	SRB	Lower 95%	Upper 95%
225	0.5039	0.4943	0.5322
226	0.4977	0.4945	0.5324
227	0.5079	0.4945	0.5324
228	0.5089	0.4945	0.5324
229	0.5113	0.4945	0.5324
230	0.5104	0.4945	0.5324
231	0.5229	0.4944	0.5324
232	0.5141	0.4944	0.5324
233	0.5178	0.4944	0.5324
234	0.5067	0.4944	0.5324
235	0.5000	0.4944	0.5324
236	0.5140	0.4944	0.5324
237	0.5211	0.4944	0.5324
238	0.5159	0.4944	0.5324
239	0.5180	0.4944	0.5324
240	0.5129	0.4944	0.5325
241	0.5009	0.4944	0.5325
242	0.5128	0.4944	0.5325
243	0.5055	0.4944	0.5325
244	0.5016	0.4944	0.5325
245	0.5291	0.4944	0.5325
246	0.5118	0.4944	0.5325
247	0.5064	0.4944	0.5325
248	0.5106	0.4944	0.5325
249	0.5167	0.4943	0.5325
250	0.5224	0.4943	0.5325
251	0.5183	0.4943	0.5325
252	0.5051	0.4943	0.5325
253	0.5127	0.4943	0.5325
254	0.5231	0.4943	0.5325
255	0.5155	0.4943	0.5325
256	0.5221	0.4943	0.5325
257	0.5063	0.4943	0.5325
258	0.5240	0.4943	0.5325
259(*)	0.4888	0.4943	0.5326
260	0.5122	0.4943	0.5326
261	0.5103	0.4943	0.5326
262	0.5063	0.4943	0.5326
263	0.5216	0.4943	0.5326
264	0.5046	0.4943	0.5326

(c) Virginia Tech Shooting in all states

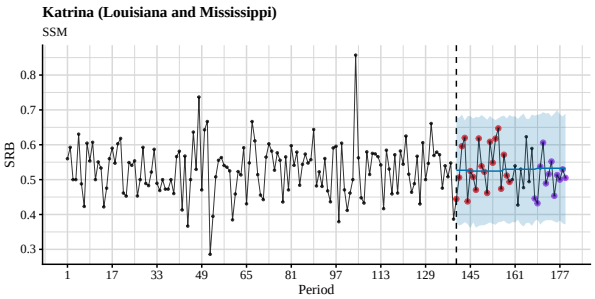
Period	SRB	Lower 95%	Upper 95%
225	0.5134	0.4567	0.5739
226	0.4908	0.4589	0.5761
227	0.5212	0.4567	0.5739
228	0.5179	0.4589	0.5761
229	0.5261	0.4567	0.5740
230	0.5322	0.4588	0.5761
231	0.5280	0.4567	0.5741
232	0.4955	0.4587	0.5761
233	0.4822	0.4567	0.5741
234	0.4647	0.4587	0.5761
235	0.5406	0.4567	0.5742
236	0.5223	0.4586	0.5761
237	0.5150	0.4567	0.5743
238	0.5421	0.4585	0.5761
239	0.5231	0.4567	0.5743
240	0.5322	0.4585	0.5761
241	0.5143	0.4567	0.5744
242	0.5197	0.4584	0.5761
243	0.5465	0.4567	0.5744
244	0.5487	0.4584	0.5761
245	0.5049	0.4567	0.5745
246	0.5105	0.4583	0.5761
247	0.5014	0.4567	0.5745
248	0.5193	0.4583	0.5761
249	0.5141	0.4567	0.5746
250	0.5068	0.4583	0.5761
251	0.5545	0.4567	0.5746
252	0.5500	0.4582	0.5761
253	0.5133	0.4567	0.5746
254	0.5172	0.4582	0.5761
255	0.5201	0.4567	0.5747
256	0.4949	0.4581	0.5761
257	0.5261	0.4567	0.5747
258	0.4946	0.4581	0.5761
259	0.5096	0.4567	0.5748
260	0.5231	0.4581	0.5761
261	0.5269	0.4568	0.5748
262	0.5155	0.4580	0.5761
263	0.5540	0.4568	0.5748
264	0.5318	0.4580	0.5761

(d) Virginia Tech Shooting in adjacent states

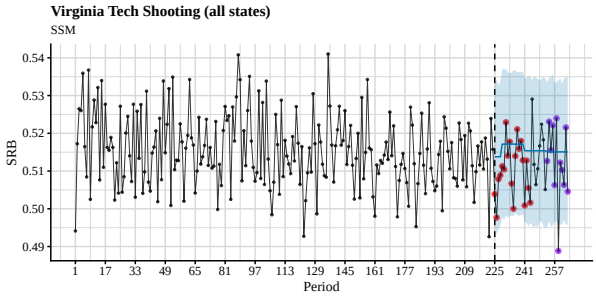
Table 4.9: Out-of-sample forecasts for the first 10 months after the intervention using SRB data grouped into 7-day periods and fitted with seasonal ARIMA models. Any period of which the observed SRB is outside of the 95% confidence level is marked by an asterisk (*).



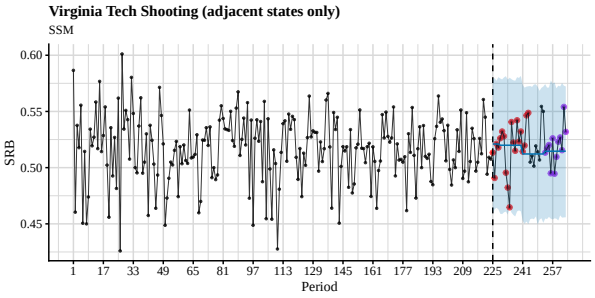
(a) Hurricane Katrina, all states



(b) Hurricane Katrina, Louisiana and Mississippi only



(c) Virginia Tech shooting, all states



(d) Virginia Tech shooting, adjacent states only

Figure 4.10: 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 5 months after the intervention are presented by red dots, whereas the observed SRBs for 7–9 months after the intervention are presented by purple dots. See also Table 4.10.

Period	SRB	Lower 95%	Upper 95%	Period	SRB	Lower 95%	Upper 95%	Period	SRB	Lower 95%	Upper 95%	Period	SRB	Lower 95%	Upper 95%
140	0.5192	0.4984	0.5388	140	0.4444	0.3793	0.6830	225	0.5039	0.4926	0.5310	225	0.5134	0.4612	0.5784
141	0.5186	0.4971	0.5400	141	0.5060	0.3685	0.6803	226	0.4977	0.4939	0.5313	226	0.4908	0.4613	0.5839
142	0.5275	0.4980	0.5384	142	0.5957	0.3726	0.6806	227	0.5079	0.4941	0.5305	227	0.5212	0.4622	0.5824
143	0.5117	0.4974	0.5399	143	0.6197	0.3769	0.6859	228	0.5089	0.4934	0.5306	228	0.5179	0.4640	0.5814
144	0.5140	0.4982	0.5394	144	0.4375	0.3680	0.6870	229	0.5113	0.4954	0.5342	229	0.5261	0.4644	0.5794
145	0.5304	0.4951	0.5371	145	0.5246	0.3642	0.6749	230	0.5104	0.4964	0.5354	230	0.5322	0.4627	0.5820
146	0.5157	0.4954	0.5392	146	0.5077	0.3789	0.6835	231	0.5229	0.4972	0.5349	231	0.5280	0.4604	0.5799
147	0.5147	0.4963	0.5382	147	0.4706	0.3716	0.6780	232	0.5141	0.4953	0.5337	232	0.4955	0.4590	0.5822
148	0.5222	0.4960	0.5387	148	0.6182	0.3691	0.6732	233	0.5178	0.4954	0.5355	233	0.4822	0.4609	0.5793
149	0.5090	0.4949	0.5386	149	0.5385	0.3728	0.6870	234	0.5067	0.4977	0.5344	234	0.4647	0.4554	0.5789
150	0.5045	0.4957	0.5378	150	0.5224	0.3742	0.6746	235	0.5000	0.4966	0.5355	235	0.5406	0.4622	0.5778
151	0.5121	0.4970	0.5383	151	0.4615	0.3784	0.6823	236	0.5140	0.4955	0.5351	236	0.5223	0.4625	0.5801
152	0.5192	0.4965	0.5390	152	0.6087	0.3665	0.6824	237	0.5211	0.4967	0.5349	237	0.5150	0.4609	0.5819
153	0.5106	0.4961	0.5392	153	0.5484	0.3744	0.6809	238	0.5159	0.4958	0.5350	238	0.5421	0.4593	0.5778
154	0.5187	0.4964	0.5390	154	0.6176	0.3691	0.6777	239	0.5180	0.4970	0.5354	239	0.5231	0.4603	0.5750
155	0.5100	0.4967	0.5386	155	0.6471	0.3808	0.6810	240	0.5129	0.4963	0.5352	240	0.5322	0.4610	0.5764
156	0.5246	0.4964	0.5369	156	0.4737	0.3655	0.6821	241	0.5009	0.4943	0.5331	241	0.5143	0.4519	0.5705
157	0.5308	0.4958	0.5378	157	0.5714	0.3753	0.6799	242	0.5128	0.4945	0.5323	242	0.5197	0.4484	0.5681
158	0.5128	0.4955	0.5366	158	0.5116	0.3833	0.6857	243	0.5055	0.4938	0.5332	243	0.5465	0.4514	0.5725
159	0.5177	0.4952	0.5375	159	0.4933	0.3718	0.6812	244	0.5016	0.4948	0.5322	244	0.5487	0.4530	0.5713
160	0.5061	0.4944	0.5392	160	0.5000	0.3801	0.6791	245	0.5291	0.4940	0.5333	245	0.5049	0.4547	0.5747
161	0.4989	0.4950	0.5385	161	0.5393	0.3729	0.6883	246	0.5118	0.4947	0.5342	246	0.5105	0.4553	0.5755
162	0.5047	0.4957	0.5377	162	0.4271	0.3790	0.6797	247	0.5064	0.4946	0.5331	247	0.5014	0.4525	0.5729
163	0.5101	0.4953	0.5385	163	0.5301	0.3742	0.6746	248	0.5106	0.4952	0.5328	248	0.5193	0.4528	0.5704
164	0.5149	0.4944	0.5382	164	0.4769	0.3739	0.6827	249	0.5167	0.4944	0.5332	249	0.5141	0.4513	0.5735
165	0.5184	0.4958	0.5383	165	0.6230	0.3781	0.6797	250	0.5224	0.4940	0.5335	250	0.5068	0.4526	0.5735
166	0.5127	0.4959	0.5368	166	0.4941	0.3730	0.6901	251	0.5183	0.4948	0.5322	251	0.5545	0.4548	0.5722
167	0.5184	0.4959	0.5369	167	0.5895	0.3765	0.6865	252	0.5051	0.4942	0.5333	252	0.5500	0.4525	0.5697
168	0.5130	0.4955	0.5370	168	0.4458	0.3765	0.6815	253	0.5127	0.4948	0.5324	253	0.5133	0.4515	0.5723
169	0.5260	0.4922	0.5352	169	0.4321	0.3759	0.6864	254	0.5231	0.4936	0.5325	254	0.5172	0.4514	0.5741
170	0.5167	0.4927	0.5339	170	0.5376	0.3806	0.6869	255	0.5155	0.4936	0.5325	255	0.5201	0.4574	0.5718
171	0.5171	0.4926	0.5364	171	0.6061	0.3742	0.6965	256	0.5221	0.4947	0.5335	256	0.4949	0.4575	0.5749
172	0.5055	0.4925	0.5340	172	0.4891	0.3844	0.6839	257	0.5063	0.4938	0.5341	257	0.5261	0.4548	0.5744
173	0.4983	0.4924	0.5348	173	0.5158	0.3761	0.6805	258	0.5240	0.4947	0.5327	258	0.4946	0.4553	0.5765
174	0.5082	0.4936	0.5351	174	0.5521	0.3795	0.6878	259(*)	0.4888	0.4944	0.5327	259	0.5096	0.4553	0.5750
175	0.5121	0.4934	0.5357	175	0.4535	0.3740	0.6792	260	0.5122	0.4946	0.5337	260	0.5231	0.4550	0.5745
176	0.5182	0.4938	0.5347	176	0.5132	0.3797	0.6959	261	0.5103	0.4944	0.5326	261	0.5269	0.4571	0.5737
177	0.5089	0.4917	0.5347	177	0.5000	0.3704	0.6787	262	0.5063	0.4936	0.5337	262	0.5155	0.4568	0.5762
178	0.5049	0.4941	0.5351	178	0.5294	0.3742	0.6895	263	0.5216	0.4938	0.5326	263	0.5540	0.4540	0.5755
179	0.5081	0.4928	0.5351	179	0.5052	0.3720	0.6802	264	0.5046	0.4938	0.5341	264	0.5318	0.4576	0.5720

(a) Hurricane Katrina in all states (b) Hurricane Katrina in Louisiana and Mississippi (c) Virginia Tech Shooting in all states (d) Virginia Tech Shooting in adjacent states

Table 4.10: Out-of-sample forecasts for the first 10 months after the intervention using SRB data grouped into 7-day periods and fitted with state-space models. Any period of which the observed SRB is outside of the 95% confidence level is marked by an asterisk (*).

4.6.5 *Swedish SRB and Meteorological Observations*

Using the data downloaded from the World Bank (`https://climateknowledgeportal.worldbank.org/download-data`), we performed a Pearson's correlation test and a Granger causality test. The p-values for the null hypotheses of the nonexistence of correlation (Student's t-test) and Granger causality (F-test) are listed in Table 4.11. We could not establish associations between the SRB and either of the meteorological observations between years 1991 and 2013.

	Temperature	Precipitation
t-test	0.156	0.765
F-test	0.269	0.228

Table 4.11: p-values for t- and F-tests on the correlation between Sweden's SRB and temperature and precipitation in Sweden

In addition, we performed logistic regression using the following:

Factor	ΔIC	SE
At risk of poverty	0.9934	3.5615
SO_2	0.8399	3.6321
NO_2	-0.5075	3.2862
Proportion foreign nationals	-1.8912	2.2786
Black smoke	-2.0551	2.7690
P80/P20	-2.2833	2.2427
Car density	-2.6865	2.4586
Population density	-2.7654	2.1618
Gini	-2.7838	2.2333
Mean income	-3.4858	2.1006
PAH	-3.7904	2.0331
VOC	-4.3064	1.8581
PM_{10}	-4.4771	1.7049
$PM_{2.5}$	-4.9524	1.4270
Median income	-5.1749	1.9721

Table 4.12: Differences in information criteria (ΔIC) and their standard errors (SE) of individual factors at the kommun (municipality) level, with random effect at the län (county) level

Table 4.13: Differences in information criteria (ΔIC) and their standard errors (SE) of individual factors at the län (county) level

Factor	ΔIC	SE
diseases of the respiratory system men	3.3012	2.5949
good health men and women	1.4435	2.4498
diseases of the circulatory system women	1.1142	3.2312
serious motor disabilities women	0.9167	3.3112
smoke and or take snuff daily men	0.8522	2.1592
unmet need for medical care men	0.7556	2.4153
motor disabilities men	0.6604	3.1667
high blood pressure women	0.2627	3.0594
smoke and or take snuff daily women	0.2378	3.1951
impaired hearing men	0.0979	2.7779
diabetes men	0.0969	3.1908
diseases of the circulatory system men	0.0838	2.7351
diseases of the respiratory system men and women	-0.1382	2.9142
serious pain total men and women	-0.1481	2.3953
serious motor disabilities men and women	-0.1740	3.0316
serious problems of anxiety worry fear men and women	-0.2266	3.5719
unmet need for medical care women	-0.2308	2.5992
poor health women	-0.3233	2.2780
smoke daily women	-0.3289	3.2831
impaired hearing women	-0.5162	2.7654
problems of anxiety worry fear men	-0.5694	2.2561
impaired vision men and women	-0.7338	2.4952
obese BMI 30 or more men	-1.0034	3.1003

Continued on next page

Factor	ΔIC	SE
serious motor disabilities men	-1.0759	2.2428
serious pain in shoulders neck women	-1.1545	2.2927
impaired vision women	-1.1893	2.4814
serious pain in hands elbows or knees men	-1.2649	3.0558
smoke daily men	-1.5168	2.7761
diseases of the musculoskeletal system and connective tissue men	-1.5226	2.4445
diseases of the skin men and women	-1.6969	3.6015
diseases of the musculoskeletal system and connective tissue men and women	-1.7217	2.6642
diabetes women	-1.8868	2.4746
severe problems from a long term illness women	-2.1122	2.7674
serious pain in shoulders neck men	-2.1671	2.8550
trouble sleeping men and women	-2.1719	2.1819
endocrine diseases men and women	-2.1860	2.4510
trouble sleeping women	-2.1972	3.5518
endocrine diseases women	-2.3205	2.5285
dentist appointments during a 12 month period men and women	-2.3418	2.8350
serious pain total men	-2.3436	4.0538
problems of anxiety worry fear men and women	-2.4328	2.5539
poor health men	-2.4561	2.3493
serious pain in back or hips men and women	-2.6307	3.2649
diseases of the circulatory system men and women	-2.7013	2.3490
high blood pressure men and women	-2.7124	2.1525
overweight or obese BMI 25 or more men and women	-2.7407	2.4722
serious pain total women	-2.7984	2.2393
diseases of the musculoskeletal system and connective tissue women	-2.8772	2.9321

Continued on next page

Factor	ΔIC	SE
diseases of the skin men	-3.0005	2.7985
dentist appointments during a 12 month period men	-3.0185	2.9935
unmet need for medical care men and women	-3.0341	1.7253
serious pain in hands, elbows, or knees women	-3.1847	3.1895
diabetes men and women	-3.2320	2.2913
serious problems of anxiety, worry, fear men	-3.2685	2.7144
serious pain in shoulders, neck men and women	-3.2852	2.7533
doctor appointments during a three month period men	-3.3952	1.9437
heart disease men and women	-3.4409	2.5030
overweight or obese BMI 25 or more men	-3.4821	2.5026
motor disabilities women	-3.5407	3.4502
obese BMI 30 or more men and women	-3.6252	2.4101
doctor appointments during a three month period men and women	-3.6527	1.6407
endocrine diseases men	-3.7796	2.4763
overweight or obese BMI 25 0 or more women	-3.9447	2.8735
severe problems from a long term illness men and women	-3.9898	2.9988
good health women	-4.0257	2.8712
heart disease men	-4.1542	2.1345
poor health men and women	-4.3023	3.0782
take snuff daily men and women	-4.3065	3.0176
problems of anxiety worry fear women	-4.3155	2.2453
serious pain in hands elbows or knees men and women	-4.3937	2.3697
impaired hearing men and women	-4.5371	2.0341
diseases of the digestive system men and women	-4.6030	2.6724
serious pain in back or hips women	-4.6210	2.2235

Continued on next page

Factor	ΔIC	SE
severe problems from a long term illness men	-4.7071	2.3077
dentist appointments during a 12 month period women	-4.7839	2.6277
take snuff daily men	-4.9910	2.1400
impaired vision men	-5.1140	2.2114
unmet need for dental care men	-5.6746	2.5992
diseases of the respiratory system women	-5.7384	2.5305
high blood pressure men	-5.7527	3.8451
serious pain in back or hips men	-5.8251	3.5053
diseases of the digestive system men	-5.9368	2.6694
unmet need for dental care men and women	-6.1628	1.9368
heart disease women	-6.2097	1.9162
take snuff daily women	-6.3476	2.0574
obese BMI 30 0 or more women	-6.4418	2.0136
motor disabilities men and women	-6.7537	1.9335
smoke and or take snuff daily men and women	-6.7639	2.4234
trouble sleeping men	-6.7890	3.4973
smoke daily men and women	-6.8021	4.0868
diseases of the nervous system and the sensory organs men	-6.8766	1.8000
diseases of the skin women	-7.0391	2.9732
diseases of the nervous system and the sensory organs men and women	-7.0903	2.1524
diseases of the digestive system women	-7.0938	2.0396
doctor appointments during a three month period women	-7.1649	4.1678
diseases of the nervous system and the sensory organs women	-7.1994	2.4814
good health men	-7.4083	1.7240
unmet need for dental care women	-8.1221	3.4038

Continued on next page

Factor	ΔIC	SE
serious problems of anxiety worry fear women	-8.7865	1.7444

4.6.6 Contingency Table Analysis

Table 4.14 is the full contingency table for testing the association between physical injury, infections, and neuropsychiatric disorders (stratified by before/during diagnosis) and SRB.

Phys_1y	Infection_1y	Neuropsych_1y	Phys_older	Infection_older	Neuropsych_older	F	M	Total
0	0	0	0	0	0	693079	733474	1426553
0	0	0	0	0	1	16835	17539	34374
0	0	0	0	1	0	65169	68766	133935
0	0	0	0	1	1	15297	16215	31512
0	0	0	1	0	0	5700	5945	11645
0	0	0	1	0	1	1265	1420	2685
0	0	0	1	1	0	5031	5250	10281
0	0	0	1	1	1	2424	2565	4989
0	0	1	0	0	0	30176	31833	62009
0	0	1	0	0	1	10905	11480	22385
0	0	1	0	1	0	5456	5688	11144
0	0	1	0	1	1	8629	9092	17721
0	0	1	1	0	0	464	537	1001
0	0	1	1	0	1	808	857	1665
0	0	1	1	1	0	510	477	987
0	0	1	1	1	1	1488	1554	3042
0	1	0	0	0	0	65857	69326	135183
0	1	0	0	0	1	3926	4229	8155
0	1	0	0	1	0	24502	25721	50223
0	1	0	0	1	1	6555	6952	13507
0	1	0	1	0	0	1230	1241	2471
0	1	0	1	0	1	352	385	737
0	1	0	1	1	0	1962	1993	3955
0	1	0	1	1	1	1077	1185	2262
0	1	1	0	0	0	9609	10417	20026
0	1	1	0	0	1	3666	3960	7626
0	1	1	0	1	0	3301	3422	6723
0	1	1	0	1	1	5243	5653	10896
0	1	1	1	0	0	174	173	347
0	1	1	1	0	1	312	316	628
0	1	1	1	1	0	271	313	584
0	1	1	1	1	1	977	1006	1983
1	0	0	0	0	0	5894	6180	12074
1	0	0	0	0	1	434	420	854
1	0	0	0	1	0	1484	1565	3049
1	0	0	0	1	1	426	469	895
1	0	0	1	0	0	245	245	490
1	0	0	1	0	1	69	70	139
1	0	0	1	1	0	227	221	448
1	0	0	1	1	1	117	138	255
1	0	1	0	0	0	1329	1454	2783
1	0	1	0	0	1	527	550	1077
1	0	1	0	1	0	258	286	544
1	0	1	0	1	1	436	482	918
1	0	1	1	0	0	28	50	78
1	0	1	1	0	1	195	203	398
1	0	1	1	1	0	28	43	71
1	0	1	1	1	1	235	256	491
1	1	0	0	0	0	2122	2204	4326
1	1	0	0	0	1	181	182	363
1	1	0	0	1	0	929	903	1832
1	1	0	0	1	1	288	305	593
1	1	0	1	0	0	69	79	148
1	1	0	1	0	1	27	31	58
1	1	0	1	1	0	143	134	277
1	1	0	1	1	1	81	87	168
1	1	1	0	0	0	650	713	1363
1	1	1	0	0	1	249	263	512
1	1	1	0	1	0	214	226	440
1	1	1	0	1	1	408	427	835
1	1	1	1	0	0	22	15	37
1	1	1	1	0	1	66	84	150
1	1	1	1	1	0	35	46	81
1	1	1	1	1	1	204	203	407
Total						1009870	1067518	2077388

Table 4.14: Contingency table of maternal diagnosis history versus the sex of livebirths

4.6.7 Dictionary of factors and their definitions

Table 4.15: List of variable names used in the main text and their corresponding definitions and units (if applicable)

Variable Name	Variable Definition	Units
A_PM10_mean_ln	particulate matter under ten micrometers in aerodynamic diameter (PM10)	ln-μg/m3
A_PM25_mean	particulate matter under 2.5 micrometers in aerodynamic diameter (PM2.5)	ln-μg/m3
A_NO2_mean_ln	nitrogen dioxide (NO2)	ln-ppb
A_SO2_mean_ln	sulfur dioxide (SO2)	ln-ppb
A_O3_mean_ln	ozone (O3)	ln-ppm
A_CO_mean_ln	carbon monoxide (CO)	ln-ppm
A_TeCA_ln	1,1,2,2-tetrachloroethane	ln-tons
A_112TCA_ln	1,1,2-trichloroethane	ln-tons
A_DBCP_ln	1,2-dibromo-3-chloropropane	ln-tons
A_TDI_ln	2,4-toluene diisocyanate	ln-tons
A_2Clacephen_ln	2-chloroacetophenone	ln-tons
A_2NP_ln	2-nitropropane	ln-tons
A_PNP_ln	4-nitrophenol	ln-tons
A_CH3CN_ln	acetonitrile	ln-tons
A_Acetophenone_ln	acetophenone	ln-tons
A_Acrolein_ln	acrolein	ln-tons
A_Acrylic_acid_ln	acrylic acid	ln-tons
A_C3H3N_ln	acrylonitrile	ln-tons
A_Sb_ln	antimony compounds	ln-tons
A_Benzidine_ln	benzidine	ln-tons
A_Benzyl_Cl_ln	benzyl chloride	ln-tons
A_Be_ln	beryllium compounds	ln-tons
A_biphenyl_ln	biphenyl	ln-tons
A_DEHP_ln	bis-2-ethylhexyl phthalate	ln-tons
A_Bromoform_ln	bromoform	ln-tons
A_Cd_ln	cadmium compounds	ln-tons
A_CS2_ln	carbon disulfide	ln-tons
A_CCl4	carbon tetrachloride	tons
A_CS_ln	carbon sulfide	ln-tons
A_Cl_ln	chlorine	ln-tons
A_C6H5Cl_ln	chlorobenzene	ln-tons
A_chloroform_ln	chloroform	ln-tons
A_Chloroprene_ln	chloroprene	ln-tons
A_Cr_ln	chromium compounds	ln-tons
A_Cresol_ln	cresol/cresylic acid	ln-tons
A_Cumene_ln	cumene	ln-tons
A_CN_ln	cyanide compounds	ln-tons
A_DBP_ln	dibutylphthalate	ln-tons
A_Diesel_ln	diesel engine emissions	ln-tons
A_DMF_ln	dimethyl formamide	ln-tons
A_Me2_phthalate_ln	dimethyl phthalates	ln-tons
A_Me2SO4_ln	dimethyl sulfate	ln-tons
A_ECH_ln	epichlorohydrin	ln-tons
A_Etacrylate_ln	ethyl acrylate	ln-tons

Continued on next page

Variable Name	Variable Definition	Units
A_EtCl_ln	ethyl chloride	ln-tons
A_EDB_ln	ethylene dibromide	ln-tons
A_EDC_ln	ethylene dichloride	ln-tons
A_EGLY_ln	ethylene glycol	ln-tons
A_EOx_ln	ethylene oxide	ln-tons
A_EdCl2_ln	ethylidene dichloride	ln-tons
A_Glycol_ethers_ln	glycol ethers	ln-tons
A_HCB_ln	hexachlorobenzene	ln-tons
A_HCBD_ln	hexachlorobutadiene	ln-tons
A_HCCPD_ln	hexachlorocyclopentadiene	ln-tons
A_Hexane_ln	hexane	ln-tons
A_N2H2_ln	hydrazine	ln-tons
A_HCl_ln	hydrochloric acid	ln-tons
A_Isophorone_ln	isophorone	ln-tons
A_Pb_ln	lead compounds	ln-tons
A_Mn_ln	manganese compounds	ln-tons
A_Hg_ln	mercury compounds	ln-tons
A_MeOH_ln	methanol	ln-tons
A_MIBK_ln	methyl isobutyl ketone	ln-tons
A_MMA_ln	methyl methacrylate	ln-tons
A_MeCl_ln	methyl chloride	ln-tons
A_Mehydrazine_ln	methylhydrazine	ln-tons
A_MTBE_ln	MTBE	ln-tons
A_nitrobenzene_ln	nitrobenzene	ln-tons
A_DMA_ln	N,N-dimethylaniline	ln-tons
A_otoluidine_ln	o-toluidine	ln-tons
A_PAHPOM_ln	PAH/POM	ln-tons
A_PCP_ln	pentachlorophenol	ln-tons
A_PH3_ln	phosphine	ln-tons
A_P_ln	phosphorus	ln-tons
A_PCBs_ln	polychlorinated biphenyls	ln-tons
A_ProCl2_ln	propylene dichloride	ln-tons
A_ProO_ln	propylene oxide	ln-tons
A_Quinoline_ln	quinoline	ln-tons
A_Se_ln	selenium compounds	ln-tons
A_Styrene_ln	styrene	ln-tons
A_Cl4C2_ln	tetrachloroethylene	ln-tons
A_Toluene_ln	toluene	ln-tons
A_C2HCl3_ln	trichloroethylene	ln-tons
A_Et3N_ln	triethylamine	ln-tons
A_VyAc_ln	vinyl acetate	ln-tons
A_VyCl_ln	vinyl chloride	ln-tons
A_11DCE_ln	vinylidene chloride	ln-tons
D303_Percent	% of stream length impaired in county	percent
SEWAGENPDESperKM	sewage permits per 1000 km of stream in county	permits/1000km
INDNPDESperKM	industrial permits per 1000 km of stream in county	permits/1000km
STORMNPDESperKM	stormwater permits per 1000 km of stream in county	permits/1000km
numDays_Close_Activity_tot	# of days closed per event in county 2000-2005	days
numDays_Cont_Activity_tot	# of days per contamination advisory event in county 2000-2005	days
numDays_Rain_Activity_tot	# of days per rain advisory event in county 2000-2005	days
Per_TotPopSS_ave	percent of population on self supply, average 2000&2005	percent

Continued on next page

Variable Name	Variable Definition	Units
Per_PSWithSW_ave	percent of public supply population which is on surface water, average 2000&2005	percent
Ca_ln_ave	calcium (Ca) precipitation weighted mean	ln mg/L
Mg_ln_ave	magnesium (Mg) precipitation weighted mean	ln mg/L
K_ln_ave	potassium (K) precipitation weighted mean	ln mg/L
Na_ln_ave	sodium (Na) precipitation weighted mean	ln mg/L
NH4_mean_ave	ammonium (NH4) precipitation weighted mean	mg/L
NO3_mean_ave	nitrate (NO3) precipitation weighted mean	mg/L
Cl_ln_ave	chloride (Cl) precipitation weighted mean	ln mg/L
SO4_mean_ave	sulfate (SO4) precipitation weighted mean	mg/L
Hg_ln_ave	total mercury (Hg) deposition	ln mg/L
AvgOfD3_ave	% of county drought – extreme (3-D4)	percent
W_As_ln	arsenic	ln mg/L
W_Ba_ln	barium	ln mg/L
W_Cd_ln	cadmium	ln mg/L
W_Cr_ln	chromium	ln mg/L
W_CN_ln	cyanide	ln mg/L
W_FL_ln	fluoride	ln mg/L
W_HG_ln	mercury (inorganic)	ln mg/L
W_NO3_ln	nitrate	ln mg/L
W_NO2_ln	nitrite	ln mg/L
W_SE_ln	selenium	ln mg/L
W_Sb_ln	antimony	ln mg/L
W_Be_ln	beryllium	ln mg/L
W_Tl_ln	thallium	ln mg/L
W_Endrin_ln	endrin	ln mg/L
W_Lindane_ln	lindane	ln mg/L
W_methoxychlor_ln	methoxychlor	ln μg/L
W_Toxaphene_ln	toxaphene	ln μg/L
W_Dalapon_ln	dalapon	ln μg/L
W_DEHA_ln	di(2-ethylhexyl)adipate (DEHA)	ln μg/L
W_Oxamyl_ln	oxamyl (Vydate)	ln μg/L
W_Simazine_ln	simazine	ln μg/L
W_DEHP_ln	Di(2-ethylhexyl) phthalate (DEHP)	ln μg/L
W_Picloram_ln	picloram	ln μg/L
W_Dinoseb_ln	dinoseb	ln μg/L
W_HCCPD_ln	hexachlorocyclopentadiene	ln μg/L
W_Carbofuran_ln	carbofuran	ln μg/L
W_atrazine_ln	atrazine	ln μg/L
W_Alachlor_ln	alachlor	ln μg/L
W_Heptachlor_ln	heptachlor	ln μg/L
W_Heptachlor_epox_ln	heptachlor epoxide	ln μg/L
W_24D_ln	2,4-D (2,4-Dichlorophenoxyacetic acid)	ln μg/L
W_HCB_ln	hexachlorobenzene	ln μg/L
W_BenzoAP_ln	benzo[a]pyrene	ln μg/L
W_PCP_ln	pentachlorophenol	ln μg/L
W_124TCIB_ln	1,2,4-trichlorobenzene	ln μg/L
W_PCB_ln	polychlorinated biphenyls (PCBs)	ln μg/L
W_DBCP_ln	1,2-dibromo-3-chloropropane (DBCP)	ln μg/L
W_EBD_ln	ethylene dibromide (EDB)	ln μg/L
W_xylenes_ln	xylenes	ln μg/L
W_Chlordane_ln	chlordane	ln μg/L

Continued on next page

Variable Name	Variable Definition	Units
W_DCM_ln	dichloromethane (methylene chloride)	ln μg/L
W_ODCB_ln	1,2-dichlorobenzene (o-dichlorobenzene)	ln μg/L
W_PDCB_ln	1,4-dichlorobenzene (p-dichlorobenzene)	ln μg/L
W_VCM_ln	vinyl chloride	ln μg/L
W_11DCE_ln	1,1-dichloroethylene	ln μg/L
W_t12DCE_ln	trans-1,2-Dichloroethylene	ln μg/L
W_EDC_ln	1,2-dichloroethane (Ethylene Dichloride)	ln μg/L
W_111trichlorane_ln	1,1,1-trichloroethane	ln μg/L
W_CCl4_ln	carbon tetrachloride	ln μg/L
W_PDC_ln	1,2-dichloropropane	ln μg/L
W_Trichlorene_ln	trichloroethylene	ln μg/L
W_112TCA_ln	1,1,2-trichloroethane	ln μg/L
W_C2Cl4_ln	tetrachloroethylene	ln μg/L
W_benzene_ln	benzene	ln μg/L
W_Cl1benz_ln	monochlorobenzene (chlorobenzene)	ln μg/L
W_Toluene_ln	toluene	ln μg/L
W_ethylbenz_ln	ethylbenzene	ln μg/L
W_styrene_ln	styrene	ln μg/L
W_DCE_ln	alpha particles	ln μg/L
W_alpha	cis-1,2-dichloroethylene	PCl/L
W_SILVEX_ln	silvex	ln μg/L
pct_harvest_acres	harvested acreage	percent
pct_irrigated_acres_ln	irrigated acreage	ln-percent
farms_per_acre_ln	farms per acre	ln-(number farms/total acres)
pct_manure_acres_ln	manure applied	ln-percent
pct_nematode_acres_ln	chemicals used to control nematodes	ln-percent
pct_disease_acres_ln	chemicals used to control disease	ln-percent
pct_defoliate_acres_ln	chemicals used to defoliate/control growth/thin fruit	ln-percent
pct_au_ln	animal units	ln-percent
herbicides_ln	herbicides	ln-pounds
fungicides_ln	fungicides	ln-pounds
insecticides_ln	insecticides	ln-pounds
mean_as_ln	arsenic	ln-ppm
mean_se_ln	selenium	ln-ppm
mean_hg_ln	mercury	ln-ppm
mean_pb_ln	lead	ln-ppm
mean_zn_ln	zinc	ln-ppm
mean_cu_ln	copper	ln-ppm
mean_na__pct_ln	sodium	ln-weighted percent
mean_mg_pct_ln	magnesium	ln-weighted percent
mean_ti_pct_ln	titanium	ln-weighted percent
mean_ca_pct_ln	calcium	ln-weighted percent
mean_fe_pct_ln	iron	ln-weighted percent
mean_al_pct	aluminum	weighted percent
mean_p_pct	phosphorus	weighted percent
facilities_rate_log	facilities per county pop	ln-rate
radon_zone	radon zone	radon category
HWYPROP	proportion of roads that are highway	miles highways / miles total roads
RYPROP	proportion of roads that are primary streets	miles primary streets / miles total roads
fatal_rate_log	traffic fatality rate	ln-rate
pct_pub_transport_log	percent of population using public transport	ln-percent

Continued on next page

Variable Name	Variable Definition	Units
rate_al_pn_gm_env_log	vice-related businesses	ln-rate
rate_ent_env_log	entertainment-related businesses	ln-rate
rate_ed_env_log	education-related businesses	ln-rate
rate_food_env_neg	negative food related businesses	rate
rate_food_env_pos_log	positive food related businesses	ln-rate
rate_hc_env_log	health care-related businesses	ln-rate
rate_rec_env_log	recreation-related businesses	ln-rate
rate_trans_env_log	transportation-related businesses	ln-rate
rate_civic_env_log	civic-related businesses	ln-rate
to_unit_rate_log	total subsidized housing units	ln-rate
pct_rent_occ	percent renter occupied	count county occupied rental units / total county units
pct_vac_units	percent vacant units	count county vacant units / total housing units
med_hh_value	median household value	dollars
med_hh_inc	median household income	dollars
pct_pers_lt_pov	percent persons less than poverty level	count county persons below poverty / county population
pct_no_eng	percent no English	count county non-English speaking / county population
pct_hs_more	percent earning greater than high school education	count county more than high school / county population
pct_unemp	percent unemployed	count county unemployed / county population
work_out_co	percent work outside county	count county work outside county / county population
med_rooms	median number rooms per house	sum county count of rooms / county housing units
pct_mt_10units_log	percent of housing with more than 10 units	ln-percent
violent_rate_log	mean number of violent crimes per capita	ln-rate
fips	FIPS code to state and county level	N/A
county_name	name of county	N/A
state	state	N/A
cat_rucc	rural-urban continuum code category	N/A
air_EQI_22July2013	non-stratified air domain index	N/A
water_EQI_22July2013	non-stratified water domain index	N/A
land_EQI_22July2013	non-stratified land domain index	N/A
sociod_EQI_22July2013	non-stratified sociodemographic domain index	N/A
built_EQI_22July2013	non-stratified built environment domain index	N/A
EQI_22July2013	non-stratified environmental quality index	N/A
RUCC1_air_EQI_22July2013	metropolitian-urbanized strata air domain index	N/A
RUCC1_water_EQI_22July2013	metropolitian-urbanized strata water domain index	N/A
RUCC1_land_EQI_22July2013	metropolitian-urbanized strata land domain index	N/A
RUCC1_sociod_EQI_22July2013	metropolitian-urbanized strata sociodemographic domain index	N/A
RUCC1_built_EQI_22July2013	metropolitian-urbanized strata built environment domain index	N/A
RUCC1_EQI_22July2013	metropolitian-urbanized strata environmental quality index	N/A
RUCC2_air_EQI_22July2013	non-metropolitian-urbanized strata air domain index	N/A
RUCC2_water_EQI_22July2013	non-metropolitian-urbanized strata water domain index	N/A
RUCC2_land_EQI_22July2013	non-metropolitian-urbanized strata land domain index	N/A
RUCC2_sociod_EQI_22July2013	non-metropolitian-urbanized strata sociodemographic domain index	N/A
RUCC2_built_EQI_22July2013	non-metropolitian-urbanized strata built environment domain index	N/A
RUCC2_EQI_22July2013	non-metropolitian-urbanized strata environmental quality index	N/A
RUCC3_air_EQI_22July2013	less-urbanized strata air domain index	N/A
RUCC3_water_EQI_22July2013	less-urbanized strata water domain index	N/A
RUCC3_land_EQI_22July2013	less-urbanized strata land domain index	N/A
RUCC3_sociod_EQI_22July2013	less-urbanized strata sociodemographic domain index	N/A
RUCC3_built_EQI_22July2013	less-urbanized strata built environment domain index	N/A
RUCC3_EQI_22July2013	less-urbanized strata environmental quality index	N/A
RUCC4_air_EQI_22July2013	rural strata air domain index	N/A

Continued on next page

Variable Name	Variable Definition	Units
RUCC4_water_EQI_22July2013	rural strata water domain index	N/A
RUCC4_land_EQI_22July2013	rural strata land domain index	N/A
RUCC4_sociod_EQI_22July2013	rural strata sociodemographic domain index	N/A
RUCC4_built_EQI_22July2013	rural strata built environment domain index	N/A
RUCC4_EQI_22July2013	rural strata environmental quality index	N/A

CHAPTER 5

INCIDENCES OF EARLY LIFE IMMUNE SYSTEM AND NEURODEVELOPMENTAL DISORDERS ARE LINKED WITH PERI- AND POSTPARTUM HEALTH FACTORS

5.1 Introduction

Understanding the aetiology of neurodevelopmental disorders (NDDs) of childhood is arguably one of the most important problems of current biomedicine (Parenti et al., 2020). Three of the most prevalent childhood NDDs are: autism spectrum disorders (ASD), attention-deficit/hyperactivity disorder (ADHD), and learning disorder (LD), which is sometimes split into reading- and mathematics disorders (Cantwell and Baker, 1991). The ASD comprises a range of severity of outwardly similar cognitive deficits disrupting a child's social interactions, her ability to read emotions, and associated with a plethora of repetitive behaviors and comorbid illnesses. ADHDThe ADHD is a milder but more prevalent condition manifesting itself in a child's impulsivity, inability to focus and hyperactivity. LD is the mildest condition among the three: an LD-affected child has normal intelligence and motivation but struggles with one or more areas of learning (such as mathematics, reading, or writing). ASD prevalence varies geographically, with nearly a hundred of recent estimates placing the mean ASD prevalence at 1%, with a full range of estimates covering a hundred times lower- and five-times higher values (Zeidan et al., 2022). Meanwhile, its heritability is estimated to be around 95% (Tick et al., 2016). ADHD together with hyperkinetic disorder (HD) is estimated to have a world prevalence of 5.29% and heritability of 80% (Polanczyk et al., 2007). LD has a prevalence of $4 \sim 9\%$ for deficits in reading and $3 \sim 7\%$ for deficits in mathematics (Moll et al., 2014). The heritability of LD is estimated to be between 40% and 70% depending on area of learning deficit (Gialluisi et al., 2021).

While a simple mathematical model-based heritability estimates suggest a major role of inherited genetic variation in predisposition to these disorders, aetiology of NDDs remains unknown. It

has been suspected that a significant contribution to disease aetiology comes from environmental insults and the concerted action of specific environmental encounters in a child's development with inherited genetic variants. One of the leading hypotheses in the space of environmental triggers of NDDs is maternal immune activation (MIA) during pregnancy 8,9: an external environmental trigger, such as maternal infection, exposure to pollutants, or medications, alters the state of the pregnant mother's immune system. Molecular signals of maternal inflammation, cytokines, has been hypothesized to leak into the developing organs of the fetus, generating persistent neuroinflammation, which derails normal maturation of the brain.

The goals of this study are (a) identify a very large mother-newborn pair cohort (IBM Watson Health, 2019; Kulaylat et al., 2019) aiming to examine the effect on child's immune and cognitive health of infectious agents affecting mother or fetus, of maternal inflammatory and autoimmune conditions, and of maternal and newborn-specific anti-infective medications; (b) probe association of specific peri- and postnatal environmental triggers with the immune health of the newborn; and (c) test associations of NDD phenotypes with environmental triggers and the state of the child's immune system. The full range of covariates that we consider in this study includes air pollution measurements (PM2.5); mother and newborn immune-system related diseases; mother and newborn anti-infective medications; Caesarean section (C-section) mode of delivery as opposed to vaginal delivery; preterm and severely preterm birth contrasted with the full-term pregnancy; abnormal and normal newborn weights at birth; age of a mother at childbirth, distinguishing teenage, typical-age, and advanced-age groups of pregnant women; year of birth as well as time of birth in year. Further more, to assess states of immune systems of a mother and her child, we recorded the counts of infections of mother and her child, distinguishing between bacterial, viral, and fungal infections. In addition, we considered sequelae of infectious diseases, defined as medical complications following infections. For example, bacterial pneumonia can be a sequela of preceding seasonal influenza and infection with intracellular parasite *Toxoplasma gondii* can have a broad range of sequelae, affecting patient's sensory systems and even inducing severe psychiatric conditions (Tamer et al., 2008;

Lykins et al., 2016).

5.2 Method

To model the risk of immune system outcomes of the newborns, represented as expected counts of the outcomes, we used negative binomial (NB) regression with an optional hurdle component. Compared to the more widely-used Poisson regression model, the NB model drops the highly restrictive assumption that the mean and variance of the distribution of counts of an outcome be equal. To model the risk of developing the NDDs of the newborns, we used logistic regression using the presence/absence of the given NDD phenotype as the outcome. For outcomes we used the following: ADHD, ASD and Learning Difficulties because the predictive reliability of using claim-based databases for identification had been previously established (Straub et al., 2021). To identify the cases, we used the ICD-9 and ICD-10 codes and algorithms in (Straub et al., 2021, Table 1).

In both classes of models we included a range of diverse risk factors for the predictors. The predictors sex, C-section and preterm/very preterm birth were coded as binary variables, while birth weight and maternal age at birth were coded as categorical variables with 3 distinct values. Furthermore, we divided $PM_{2.5}$ into quintiles. Finally, for immune-system related health events and anti-infective prescriptions we used an indicator variable for each sub-variable due to the small numbers of subjects with non-zero values. To preserve causality, we used a cut-off period of 6 months for the immune system submodels whereby the diagnosis and prescription predictors were counted only before the cutoff, and outcomes were counted only after the cutoff. See Table 5.1 for more details.

Abbreviation	code	description
S	`sex`	Newborn sex
C	`csec`	Caesarean section
P	`prem`	Gestation: preterm (32 ~ 37 wk), very preterm (< 32 wk)
W	`weight`	Birth weight: low ($\leqslant$ 2500 g), high ($\geqslant$ 4500 g)
A	`age`	Mother's age at birth: teenage (13–19), advanced (35–45)
M	`PM25`	Particulate Matter 2.5 μm ($PM_{2.5}$)
D	`DX`	Immune-system related health events (infections, sequelae, immune disorders)
R	`RX`	anti-infective prescriptions (antibacterials, antimycotics, antiparasitics)
H	`hurd`	Hurdle term
TD	-	Day of birth in year: daily resolution (P-spline smooths)
TM	-	Day of birth in year: monthly resolution
TS	-	Day of birth in year: seasonal resolution

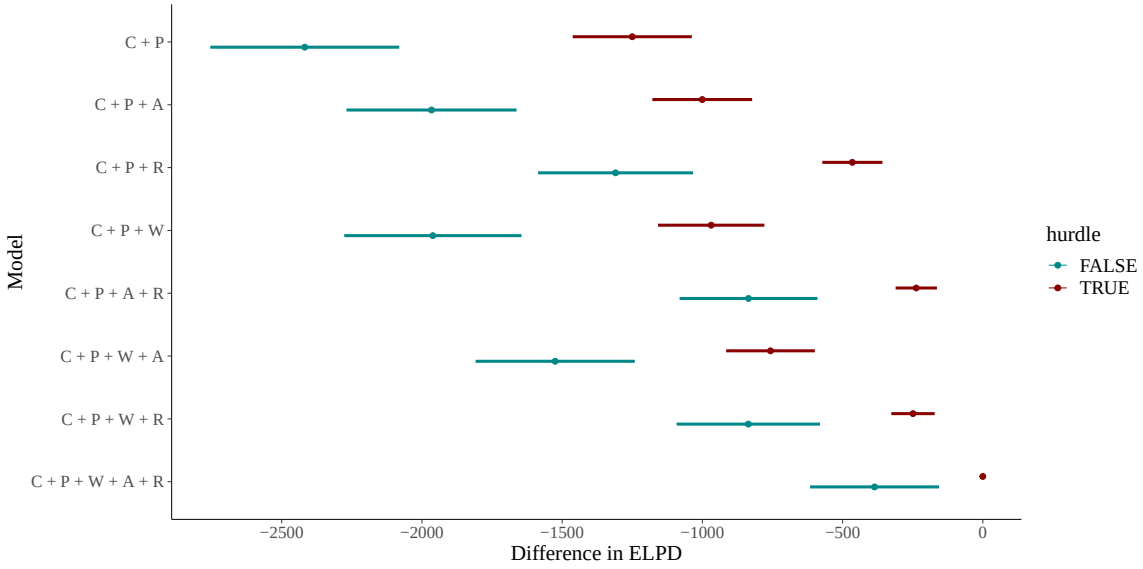
Table 5.1: Abbreviations for model predictors and other terms.

5.3 Results

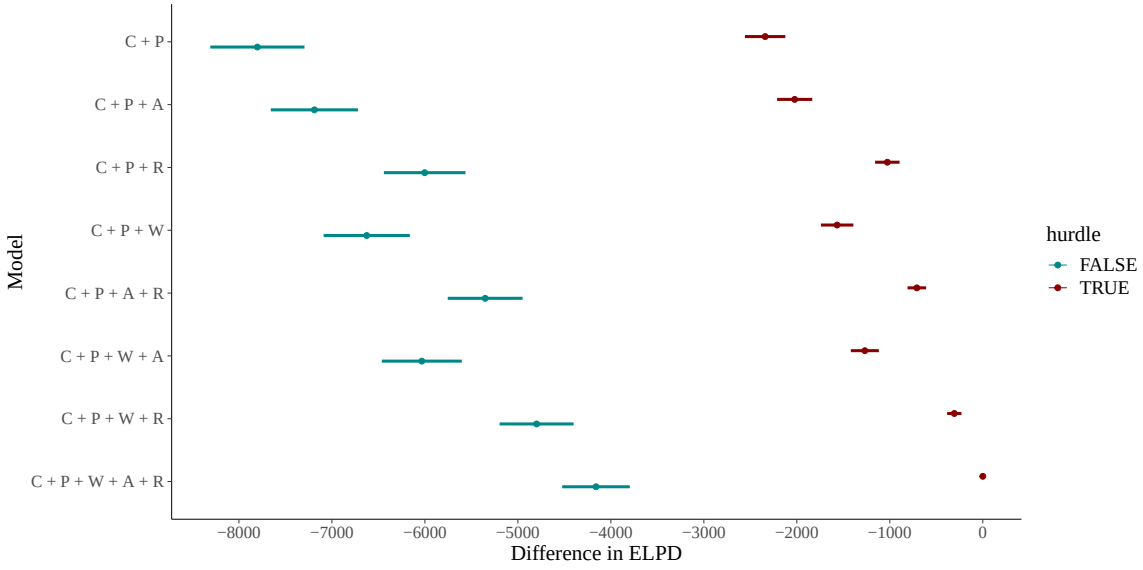
5.3.1 Model comparison

Figures 5.1 and 5.2 show the difference in ELPD (Δ_{ELPD}) w.r.t. the model with the smallest information criterion along with the corresponding 95% credible intervals for immune systems and NDD models, respectively. For the immune system submodels, the full model was the optimal model for all of the outcomes studied except miscellaneous infections, for which the optimal model was the full model without the hurdle term. Moreover, LOO diagnostics indicated that models were poorly specified while the number of problematic data points are high (> 10). In other words the models, including the priors and the likelihood, were not good at recovering the data. Using 10-fold cross validation, we found that among the NDD submodels, in the case of ADHD the full model ($C + P + W + A + M + D + R$) was the sole best model according to estimated Δ_{ELPD}, whereas for ASD and learning difficulties more than one model was optimal. By applying the Bayesian Occam's Razor Myung and Pitt (1997), we picked the most parsimonious model among the ones that were not significantly different from the one with smallest LOOIC: for ASD, $C+P+W+A+M$ and for Learning Difficulties, $C+P+M$. For all NDD models we have chosen to use the smoothing splines to represent the effect of time of birth in the year.

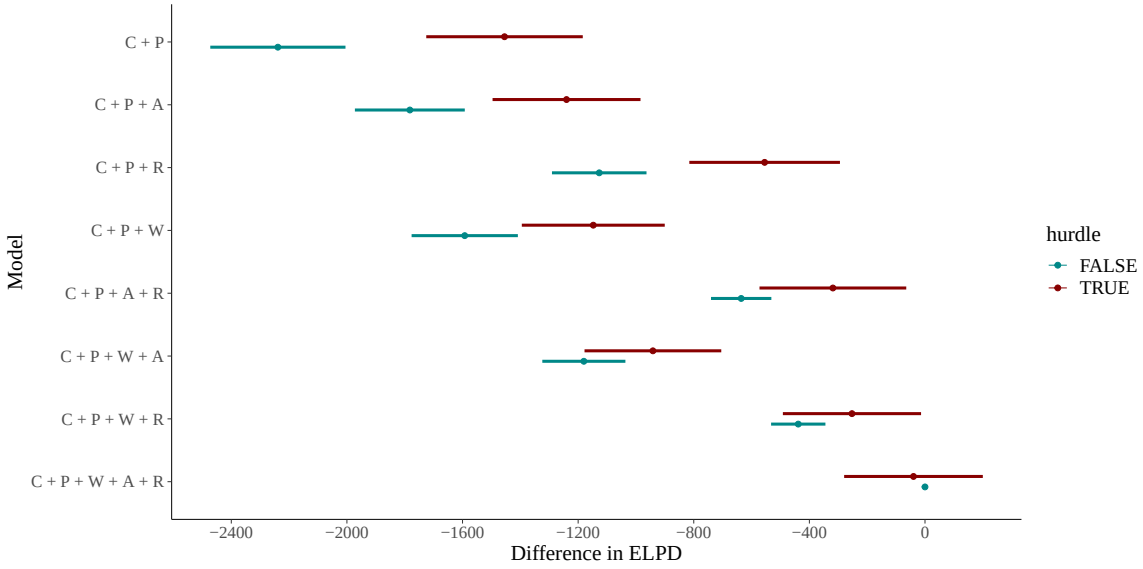
(a) Bacterial Infections



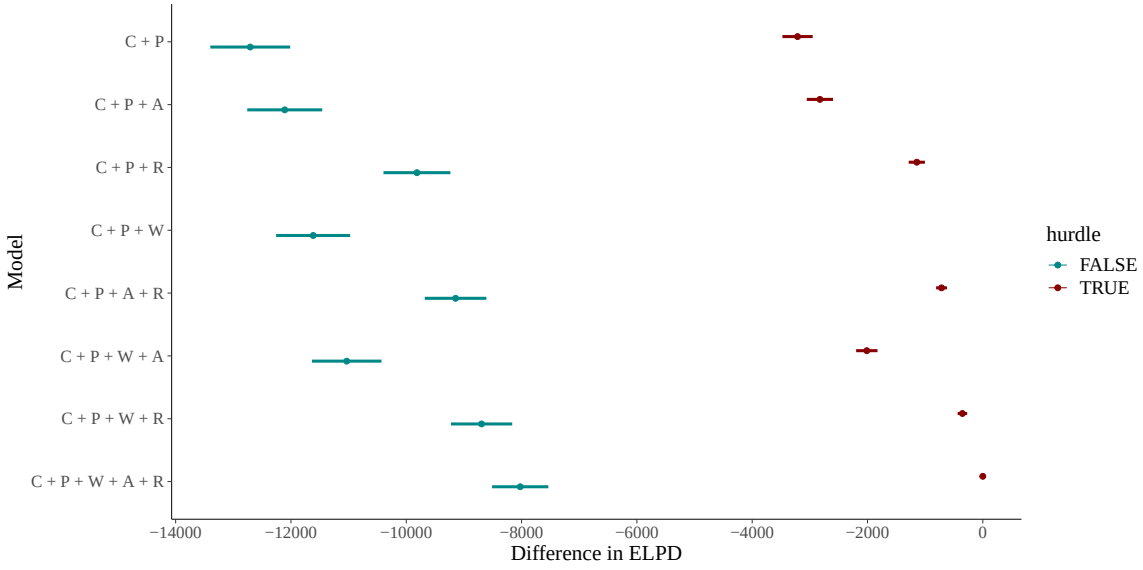
(b) Viral Infections



(c) Miscellaneous Infections



(d) Sequelae Infections



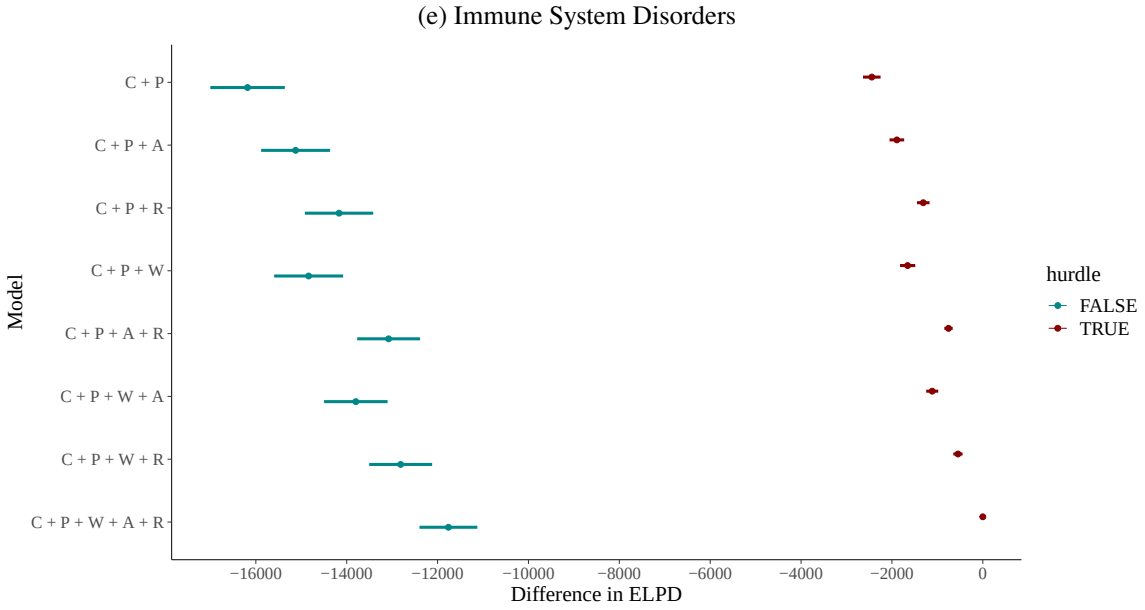
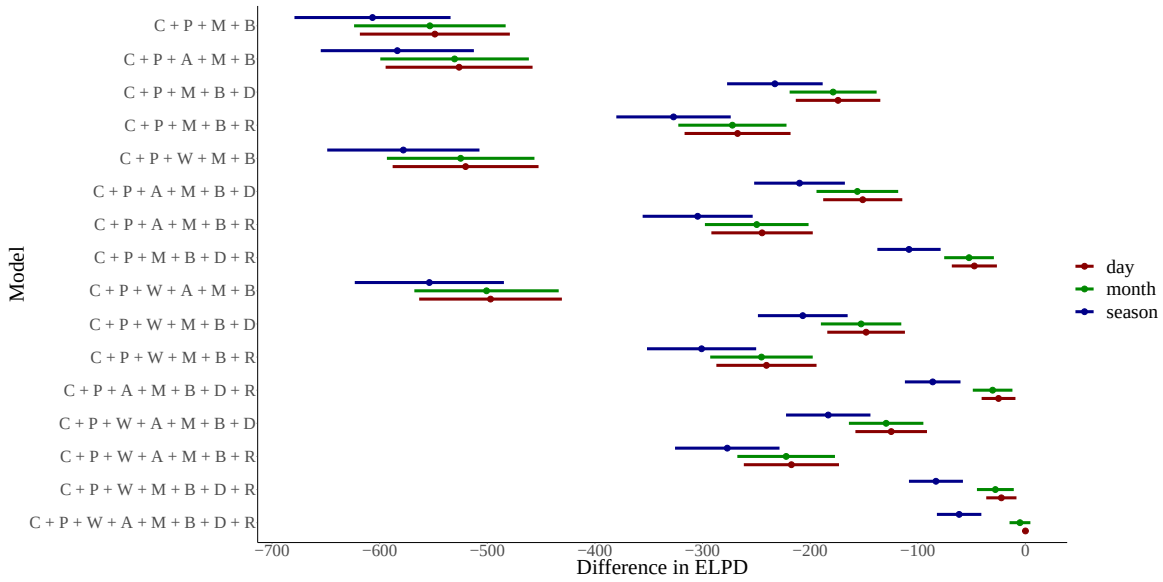
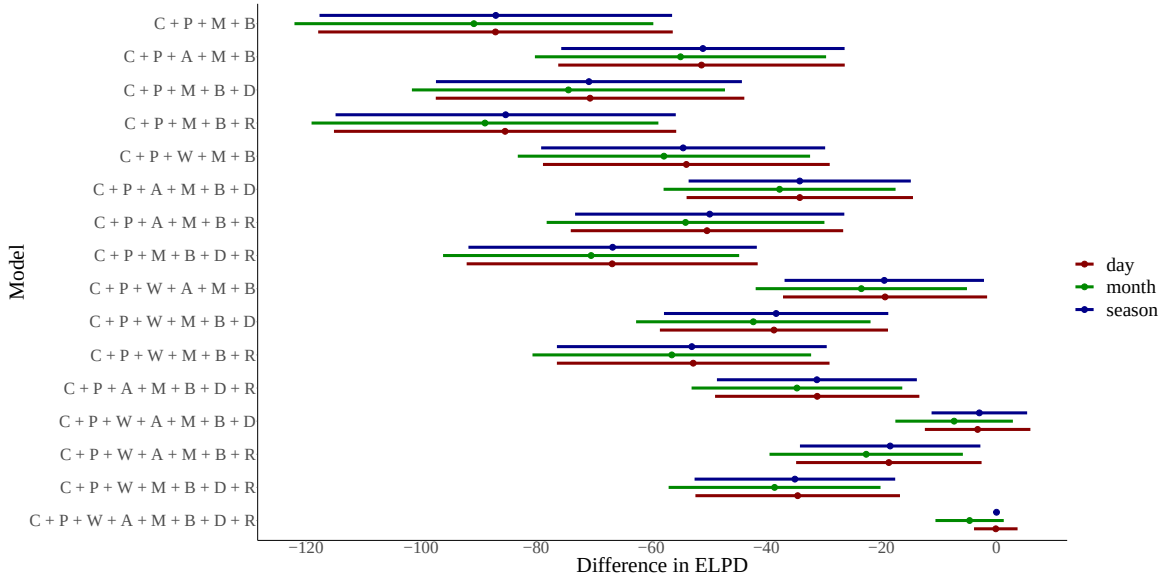


Figure 5.1: Immune systems model comparison with 10-fold cross validation: dots represent the estimates for Δ_{ELPD} and intervals represent the corresponding 95% credible interval under asymptotic normal distribution. On the y-axis the model terms are coded according to Table 5.1.

(a) ADHD


C + P + M + B
C + P + A + M + B
C + P + M + B + D
C + P + M + B + R
C + P + W + M + B
C + P + A + M + B + D
C + P + A + M + B + R
C + P + M + B + D + R
C + P + W + A + M + B
C + P + W + M + B + D
C + P + W + M + B + R
C + P + A + M + B + D + R
C + P + W + A + M + B + D
C + P + W + A + M + B + R
C + P + W + M + B + D + R
C + P + W + A + M + B + D + R
Model
−700
−600
−500
−400
−300
−200
−100
0
Difference in ELPD
day
month
season

(b) ASD


C + P + M + B
C + P + A + M + B
C + P + M + B + D
C + P + M + B + R
C + P + W + M + B
C + P + A + M + B + D
C + P + A + M + B + R
C + P + M + B + D + R
C + P + W + A + M + B
C + P + W + M + B + D
C + P + W + M + B + R
C + P + A + M + B + D + R
C + P + W + A + M + B + D
C + P + W + A + M + B + R
C + P + W + M + B + D + R
C + P + W + A + M + B + D + R
Model
−120
−100
−80
−60
−40
−20
0
Difference in ELPD
day
month
season

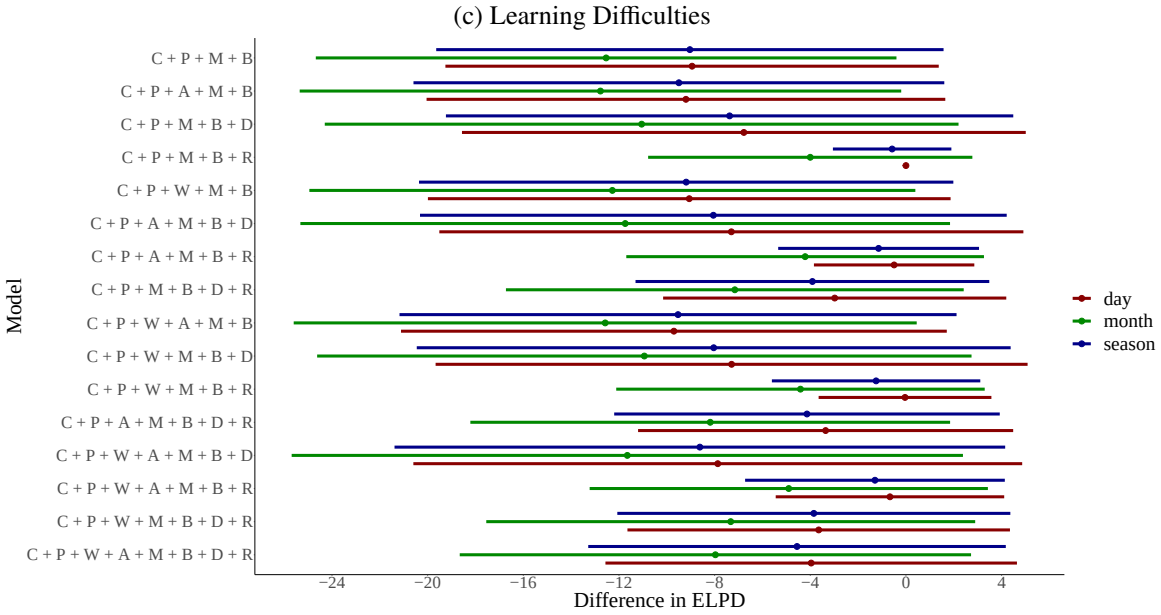


Figure 5.2: NDD model comparison with PSIS-LOO: dots represent the estimates for Δ_{ELPD} and intervals, the corresponding 95% credible interval under asymptotic normal distribution. On the *y*-axis the model terms are coded according to Table 5.1. Notice that for ASD and Learning Difficulties there was more than one model was the best.

5.3.2 *Parameter Estimation*

Risk Factors	Bacterial Infections	Immune Disorders	Miscellaneous Infections	Sequelae Infections	Viral Infections
Mother Infections: Bacterial	**39.0 (29.3, 49.3)**	**4.6 (2.7, 6.5)**	**20.9 (17.1, 24.8)**	**16.9 (13.5, 20.4)**	**10.3 (6.5, 14.1)**
Mother Infections: Viral	4.4 (-2.1, 11.4)	-1.2 (-2.8, 0.5)	**-11.6 (-14.2, -8.8)**	**8.5 (5.7, 11.3)**	**27.0 (23.6, 30.8)**
Mother Infections: Miscellaneous	**16.8 (9.8, 24.0)**	**1.7 (0.1, 3.2)**	**11.3 (8.4, 14.4)**	**8.7 (6.2, 11.3)**	**7.7 (4.8, 10.7)**
Mother Infections: Sequelae	**18.0 (10.3, 26.2)**	**5.0 (3.2, 6.9)**	-2.4 (-5.4, 0.6)	**23.8 (20.9, 27.1)**	**12.1 (8.8, 15.7)**
Mother Inflammations	-2.9 (-50.2, 92.7)	9.8 (-7.1, 29.7)	-24.3 (-45.9, 4.4)	1.3 (-22.2, 31.4)	17.0 (-13.9, 58.3)
Mother Immune Disorders	**-4.2 (-8.1, -0.2)**	**22.1 (20.7, 23.5)**	**6.6 (4.5, 8.7)**	**13.1 (11.3, 15.0)**	**9.4 (7.5, 11.5)**
Mother RX: Antibacterials	**-10.0 (-13.7, -6.3)**	**5.0 (3.8, 6.3)**	**-10.0 (-11.9, -8.1)**	**-10.8 (-12.3, -9.3)**	**4.6 (2.6, 6.8)**
Mother RX: Antimycotics	5.8 (-0.5, 12.4)	**5.8 (4.2, 7.5)**	**4.1 (1.1, 6.9)**	2.0 (-0.5, 4.6)	**8.7 (5.8, 11.8)**
Mother RX: Antiparasitics	**19.9 (10.6, 30.4)**	**-4.5 (-6.5, -2.4)**	**9.3 (5.2, 13.3)**	2.8 (-0.3, 6.2)	0.9 (-3.0, 4.9)
Newborn RX: Antibacterials	**33.0 (27.8, 38.7)**	**31.7 (30.2, 33.3)**	**20.7 (18.2, 23.2)**	**17.7 (15.7, 19.7)**	**18.7 (16.5, 20.9)**
Newborn RX: Antimycotics	**66.5 (48.8, 86.0)**	**25.7 (22.2, 29.5)**	**32.9 (26.4, 39.8)**	**25.7 (20.0, 31.6)**	**29.6 (22.6, 36.5)**
Newborn RX: Antiparasitics	**258.5 (117.2, 516.6)**	13.3 (-0.5, 29.9)	**30.0 (3.0, 62.2)**	**50.9 (20.4, 86.3)**	3.7 (-19.8, 32.4)
Newborn Sex: Female	**-9.6 (-12.9, -6.2)**	**-17.1 (-18.0, -16.1)**	**9.7 (7.7, 11.8)**	**-5.1 (-6.5, -3.6)**	0.3 (-1.4, 2.1)
Newborn Caeserean Section	**13.9 (9.6, 18.5)**	**11.1 (9.8, 12.3)**	**8.3 (6.1, 10.5)**	**14.0 (12.2, 15.9)**	**3.5 (1.7, 5.5)**
Birth Condition: Premature	**25.8 (10.1, 43.8)**	1.3 (-2.1, 4.9)	**-7.8 (-13.8, -1.7)**	**23.2 (17.0, 30.0)**	**9.7 (2.8, 17.0)**
Birth Condition: Very Premature	**173.1 (127.9, 230.5)**	**25.0 (19.1, 31.3)**	**13.9 (4.6, 24.2)**	**44.2 (34.0, 55.3)**	**15.3 (5.2, 27.3)**
Birth Condition: Caeserean Section × Premature	-14.1 (-27.8, 1.5)	0.0 (-4.4, 4.5)	**13.3 (4.7, 23.1)**	**-6.5 (-12.7, -0.6)**	2.2 (-5.8, 10.5)
Birth Condition: Caeserean Section × Very Premature	14.5 (-8.1, 42.3)	**7.2 (1.1, 13.5)**	**25.9 (13.6, 39.0)**	8.1 (-1.0, 17.6)	3.8 (-7.0, 15.9)
Birth Condition: Low birth weight	**70.8 (59.9, 82.1)**	**16.3 (14.2, 18.5)**	**5.2 (1.8, 8.7)**	**19.4 (16.2, 22.6)**	**7.4 (4.0, 11.0)**
Birth Condition: High birth weight	**41.8 (9.6, 84.2)**	-1.4 (-7.3, 5.2)	-11.2 (-21.4, 0.1)	**24.4 (12.6, 37.3)**	9.5 (-3.0, 23.6)
Birth Condition: Teenage mother	**87.4 (35.6, 159.1)**	**-28.7 (-34.4, -22.7)**	**63.1 (43.5, 84.9)**	6.3 (-8.7, 23.8)	**-24.6 (-39.1, -6.3)**
Birth Condition: Advanced-age mother	**-5.0 (-9.0, -1.0)**	**-3.9 (-5.1, -2.7)**	**-13.2 (-14.9, -11.4)**	**5.0 (3.2, 6.8)**	**3.9 (2.0, 5.8)**

Table 5.2: Parameter estimation (in percentage) for the optimal Immune System model for each outcome in terms of odds ratio (see §5.3.1 for model comparison): Posterior median and 95% credible interval of OR; entries in **boldface** indicates the corresponding credible intervals do not include 0. See also Figure 5.3.

We now report the results from the most optimal obtained by model comparison (cf. §5.3.1). Results from the immune systems models were reported as relative increase in expected event counts (hereafter *relative risk*, RR), whereas results from the neurodevelopmental disorders were in terms of odds ratio (OR) w.r.t. the baseline. Table 5.2 and Figure 5.3 show the posterior median and the 95% credible interval of the parameters under the optimal model for the immune systems outcomes, and Table 5.3 and Figures 5.4, 5.5 and 5.6 show the same for the NDDs.

Immune System Submodels

First, a wide range of anti-infective prescriptions given to newborns during the first six months of life were associated with altered risks of immune system outcomes for the newborn. Newborn medications were associated with marked risk increases for immune system-related diseases. In particular, the presence of antiparasitic drugs was associated with a 258.5 percent and a 50.9 percent increase in risk of bacterial and sequelae infections, respectively. For antibacterial medications,

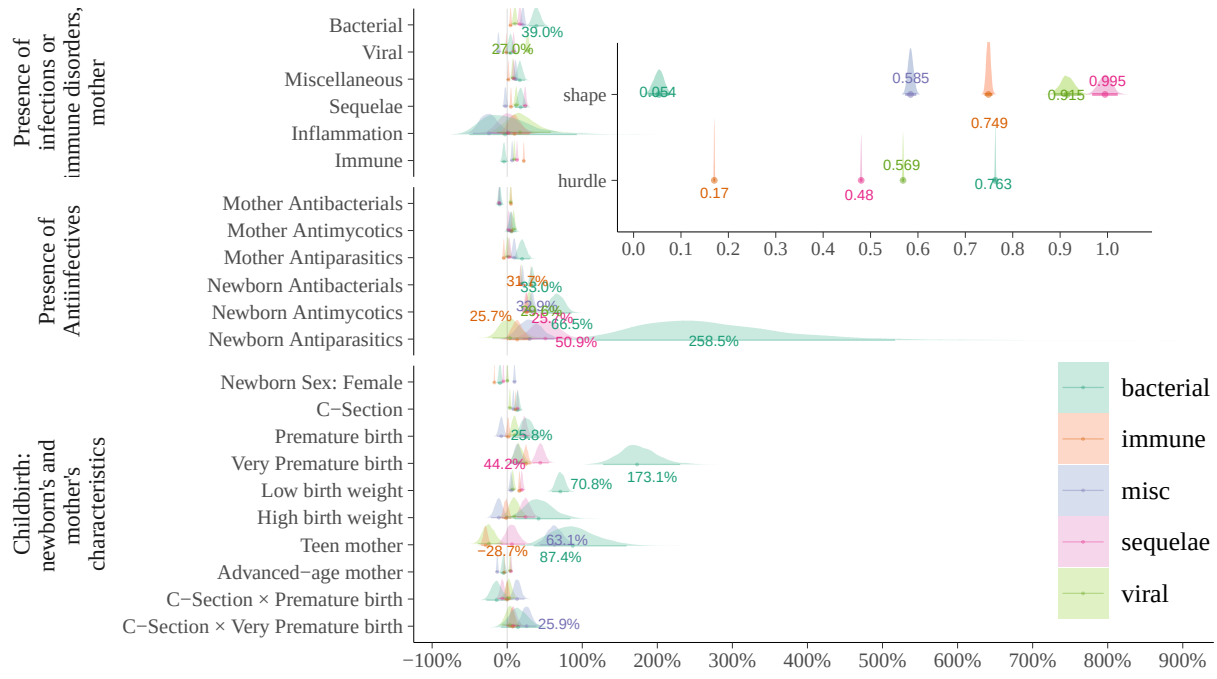


Figure 5.3: Posterior density estimation of relative risk of newborn immune system disorders for childbirth-related characteristics (S, C, P, W, A). Numerical annotations, shown only for effect sizes with absolute value greater than or equal to 25%, represent posterior medians, corresponding to the dots under the slabs. Both the slabs and the interval underneath represent 95% credible intervals. `shape` is the shape parameter of the gamma distribution underlying the negative binomial model, whereas `hurdle` is the estimated hurdle probability. Viral infection does not have an estimated `hurdle` value because the hurdle model was not the optimal model.

Risk Factors	ADHD	Autism	Learning difficulties
midrule Newborn Sex: Female	**0.39 (0.377, 0.404)**	**0.253 (0.236, 0.271)**	**0.532 (0.462, 0.613)**
Newborn Caeserean Section	**1.161 (1.125, 1.2)**	**1.187 (1.12, 1.259)**	1.091 (0.934, 1.266)
Birth Condition: Premature	1.102 (0.948, 1.274)	0.907 (0.701, 1.166)	1.625 (0.901, 2.727)
Birth Condition: Very Premature	**1.498 (1.232, 1.798)**	1.164 (0.823, 1.58)	1.981 (0.934, 3.762)
Caeserean Section × Premature	1.152 (0.966, 1.381)	1.265 (0.939, 1.709)	0.946 (0.458, 2.065)
Caeserean Section × Very Premature	1.113 (0.896, 1.382)	**1.432 (1.022, 2.08)**	1.173 (0.514, 2.836)
Birth Condition: Low birth weight	**1.288 (1.206, 1.375)**	**1.567 (1.401, 1.743)**	-
Birth Condition: High birth weight	0.858 (0.633, 1.168)	**1.749 (1.151, 2.476)**	-
Birth Condition: Teenage mother	**1.898 (1.219, 2.819)**	1.315 (0.593, 2.603)	-
Birth Condition: Advanced-age mother	**0.891 (0.859, 0.924)**	**1.306 (1.229, 1.387)**	-
PM25 quintile 2	**1.172 (1.107, 1.242)**	1.018 (0.93, 1.114)	1.117 (0.873, 1.421)
PM25 quintile 3	**1.248 (1.179, 1.32)**	1.021 (0.938, 1.117)	1.071 (0.851, 1.352)
PM25 quintile 4	**1.215 (1.148, 1.287)**	0.983 (0.897, 1.079)	0.859 (0.67, 1.108)
PM25 quintile 5	**1.258 (1.189, 1.333)**	1.072 (0.972, 1.178)	0.949 (0.746, 1.205)
Birth year: 2003	**0.427 (0.395, 0.462)**	**0.533 (0.456, 0.618)**	**0.127 (0.086, 0.186)**
Birth year: 2004	**0.547 (0.512, 0.586)**	**0.684 (0.6, 0.776)**	**0.264 (0.197, 0.35)**
Birth year: 2005	**0.653 (0.617, 0.695)**	**0.777 (0.686, 0.877)**	**0.426 (0.326, 0.551)**
Birth year: 2006	**0.85 (0.803, 0.9)**	0.917 (0.821, 1.029)	**0.59 (0.461, 0.752)**
Birth year: 2008	**1.11 (1.049, 1.178)**	**1.175 (1.055, 1.312)**	1.192 (0.925, 1.511)
Birth year: 2009	**1.234 (1.161, 1.315)**	**1.51 (1.356, 1.675)**	1.179 (0.912, 1.535)
Birth year: 2010	**1.087 (1.016, 1.162)**	**1.784 (1.603, 1.985)**	1.135 (0.84, 1.538)
Offset	**1.533 (1.523, 1.542)**	**1.314 (1.304, 1.325)**	**1.639 (1.588, 1.693)**
Mother Infections: Bacterial	**1.208 (1.121, 1.301)**	1.122 (0.973, 1.286)	-
Mother Infections: Viral	1.04 (0.973, 1.114)	0.973 (0.852, 1.101)	-
Mother Infections: Miscellaneous	1.053 (0.991, 1.121)	1.049 (0.944, 1.171)	-
Mother Infections: Sequelae	**1.109 (1.033, 1.187)**	1.087 (0.958, 1.228)	-
Mother Inflammations	0.98 (0.406, 1.912)	1.307 (0.311, 3.413)	-
Mother Immune Disorders	**1.24 (1.194, 1.285)**	1.04 (0.985, 1.102)	-
Newborn Infections: Bacterial	**1.369 (1.26, 1.498)**	1.076 (0.909, 1.272)	-
Newborn Infections: Viral	**1.232 (1.109, 1.364)**	0.903 (0.72, 1.114)	-
Newborn Infections: Miscellaneous	**1.18 (1.102, 1.261)**	**1.271 (1.132, 1.426)**	-
Newborn Infections: Sequelae	**1.232 (1.168, 1.297)**	1.069 (0.967, 1.181)	-
Newborn Inflammations	0.682 (0.017, 13.335)	0.793 (0.018, 17.829)	-
Newborn Immune Disorders	**1.154 (1.113, 1.199)**	**1.163 (1.087, 1.24)**	-
Mother RX: Antibacterial	**1.147 (1.106, 1.19)**	-	-
Mother RX: Antimycotic	**1.169 (1.103, 1.239)**	-	-
Mother RX: Antiparasitic	1.042 (0.957, 1.132)	-	-
Newborn RX: Antibacterial	**1.231 (1.188, 1.277)**	-	-
Newborn RX: Antimycotic	**1.19 (1.058, 1.338)**	-	-
Newborn RX: Antiparasitic	1.08 (0.569, 1.834)	-	-

Table 5.3: Parameter estimation for the optimal NDD model for each outcome in terms of odds ratio (see §5.3.1 for model comparison): Posterior median and 95% credible interval of OR; entries in **boldface** indicates the corresponding credible intervals do not include 0. A "-" indicates that the corresponding entry is not part of the optimal model for the given outcome phenotype. The birth year predictors were coded as a categorical variable with the year 2007 chosen as the reference level. See also Figures 5.4, 5.5 and 5.6.

the risk increases were 33.0 percent, 20.7 percent, 18.7 percent and 17.7 percent for bacterial, miscellaneous, viral and sequelae infections, respectively; risk of immune disorders increased by 31.7 percent. Finally, antimycotic medications predicted risk increases of 66.5 percent, 32.9 percent, 29.6 percent, and 25.7 percent for bacterial, miscellaneous, viral and sequelae infections, respectively, whereas for immune disorders the risk increase was 25.7 percent.

Second, medications administered to the mother turned out to have rather diverse effects on her newborn's immune health. Somewhat surprisingly, antibacterial medications appeared to be associated with lower risks for newborn's bacterial- (-10.0 percent), sequelae- (-10.8) and miscellaneous (-10.0 percent) infections. The same medications were associated with risk increases for immune disorders (5.0 percent) and viral infections (4.6 percent). Antiparasitic medications were associated with risk increases of 19.9 percent and 9.3 percent for bacterial and miscellaneous infections, respectively, while the risk decreased by 4.5 percent for immune disorders. Finally, for antimycotics, the only significant result was miscellaneous infections (4.1 percent), viral infections (8.7 percent) and immune disorders (5.8 percent).

Third, our results suggested that preterm birth, teenage pregnancy, and C-section delivery were the most consequential for newborn immune health. Female newborns had a lower risk of immune-related diseases than male, with an exception for miscellaneous infections, whereas teenage births had variable-sign effects on the outcomes:

- As expected, female newborns appear to be better protected against diseases with immune system aetiology. We estimated a -9.6 percent risk change for bacterial infections in female newborns, a -5.1 percent for sequelae infections and a -17.1 percent for immune disorders. Only for miscellaneous infections did we observe a positive risk change for female newborns (9.7 percent).

- Preterm and very preterm birth had very strong associations with a child's subsequent bacterial infections, with the disease risk increased by 173.1 percent (very preterm) and 25.8 percent (preterm). Moreover, very preterm birth was associated with 13.9 percent, 44.2 percent,

and 15.3 percent risk increases of the child's miscellaneous, sequelae, and viral infections, respectively. A child's very preterm birth is associated with a 25.0 percent risk increase of immune disorders. Interestingly, for miscellaneous infections preterm birth was associated with lower risk (-7.8 percent) but very preterm 13.9 percent.

- Childbirth via C-section was associated with 13.9 percent, 8.3 percent, 14.0 percent and 3.5 percent risk increases for newborn bacterial, miscellaneous, sequelae and viral infections, respectively. The effect size is not statistically different from zero for viral and miscellaneous infections. The risk increase associated with C-section is 3.8 percent for future immune disorders. For the interaction between C-section and preterm birth we found, among other things, risk increases in both interactions (C-section $\times$ preterm/very preterm) for miscellaneous infections (13.9 percent and 25.9 percent, respectively) even though the effect of preterm birth tended towards lower RR (see above). Similarly, for sequelae infections the interaction effect between C-section and premature birth was protective (-6.5) whereas C-section (14.0), preterm birth (23.2) and very preterm birth were all associated with higher risks.

- Children with low birth weights (less than or equal to 2,500 g) had risk increases of 70.8 percent, 27.0 percent, 7.4 percent and 5.2 percent for bacterial, sequelae, viral and miscellaneous infections. The risk increase associated with low birth weights was 16.3 percent for immune disorders. For higher birth weight (more than or equal to 4,500 g), bacterial infections and sequelae infections had elevated RR of 41.8 percent and 24.4 percent, respectively.

- Teenage pregnancy represents the most diverse effect-sign disease outcome, with a risk change of 87.4 percent and 63.1 percent for bacterial, miscellaneous and viral infections, respectively, but -28.7 percent and -24.6 percent for immune disorders and viral infections, respectively.

Finally, a pregnant woman's history of infections and immune disorder predicts predominantly significant changes in her newborn's disease risk. If a pregnant mother ever had bacterial infections, our study shows increases in risk of her newborn's bacterial infections, miscellaneous infections,

sequelae infections and viral infections of 39.0 percent, 20.9 percent, 16.9 percent and 10.43 percent, respectively. Presence of maternal sequelae infections during pregnancy predicts an increase in a newborn's risk of the same kind of infections by 23.8 percent, and an increase of 18.0 in the risk of bacterial infections. Finally, the presence of maternal viral infections is associated with a 27.0 percent increase in the risk of the same infections in the newborn. By contrast, the presence of maternal viral infections is associated with a -11.6 percent change in risk for newborn's miscellaneous infections.

NDD Submodels

Figure 5.4 shows the posterior density estimation for childbirth-related characteristics, sex, anti-infective prescriptions and maternal immune system-related events. Premature birth, C-section and female sex were the common risk factors among the outcomes. Compared to male (baseline), the newborn being female had ORs of 0.390, 0.253 and 0.532 for ADHD, ASD and Learning Difficulties, respectively, indicating drastic decreases in risks. C-section births turned out to be a significant risk factor for ADHD and ASD, with ORs of 1.161 and 1.187, respectively. By contrast, very premature births were associated with significant risk increases for ADHD only (OR = 1.498), but the interaction between C-section and very premature births had an OR of 1.432 for ASD only.

Included in the respective optimal model of ADHD and ASD but not in that of Learning Difficulties were birth weight and abnormal maternal ages. Both risk factors were coded as three-valued categorical variables with normal values as the baseline. For low birth weight, we found elevated ORs of 1.288 and 1.567 for ADHD and ASD, respectively, while advanced maternal ages predicted higher risks for ASD (OR: 1.749).

For ADHD, mother and newborn diagnosis (D) and prescription (R) information included in the optimal model, while for ASD only newborn diagnosis (D) was included (neither for Learning difficulties). In particular, bacterial infection-related factors were all associated with increased risks: maternal infection with an OR of 1.369, maternal antibacterials with an OR of 1.147, and

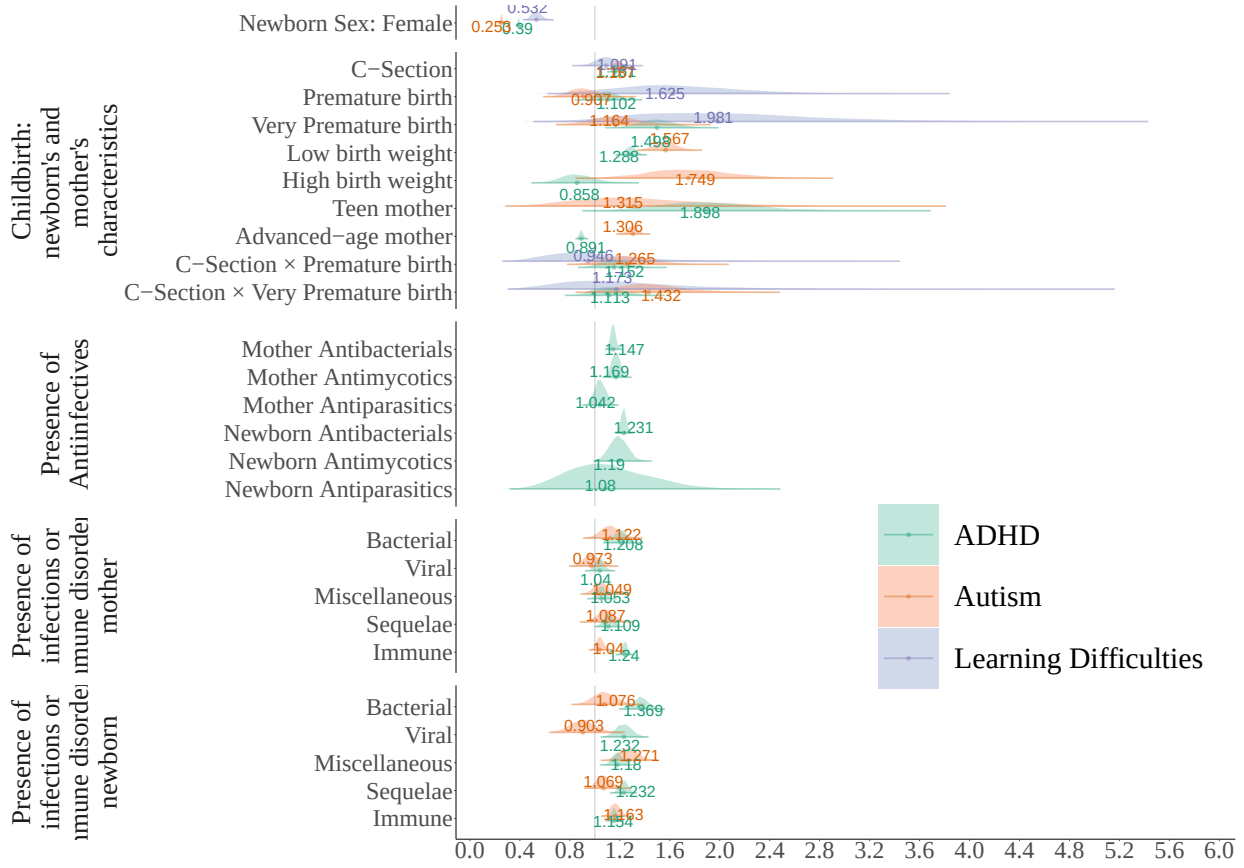


Figure 5.4: Posterior density estimation of OR for childbirth-related characteristics (S, C, P, W, A). Numerical annotations on the figure represent posterior medians, corresponding to the dots under the slabs. Both the slabs and the interval underneath represent 95% credible intervals.

newborn antibacterials with an OR of 1.231. The other type of maternal infections that was associated with an elevated risk was sequelae, with an OR of 1.109. As for other types of anti-infective prescriptions, both maternal and newborn anti-mycotics were associated with elevated ORs: 1.169 and 1.190 for mothers and newborns, respectively. Finally, presence of maternal immune disorders predicted higher newborn risks of ADHD (OR: 1.154) and for ASD (OR: 1.163).

Figure 5.5 shows the the posterior density estimation for $PM_{2.5}$, which was a significant risk factor for ADHD only. The lowest quintile, i.e. with the lowest numerical $PM_{2.5}$ readings, was used as the baseline in calculating the ORs for the other quntiles, which were 1.172, 1.248, 1.215 and 1.258 for the second to the fifth, respectively. Figure 5.6 shows the posterior density estimation for year of birth plus number of years under risk used as offset. The year 2007 was used as the baseline.

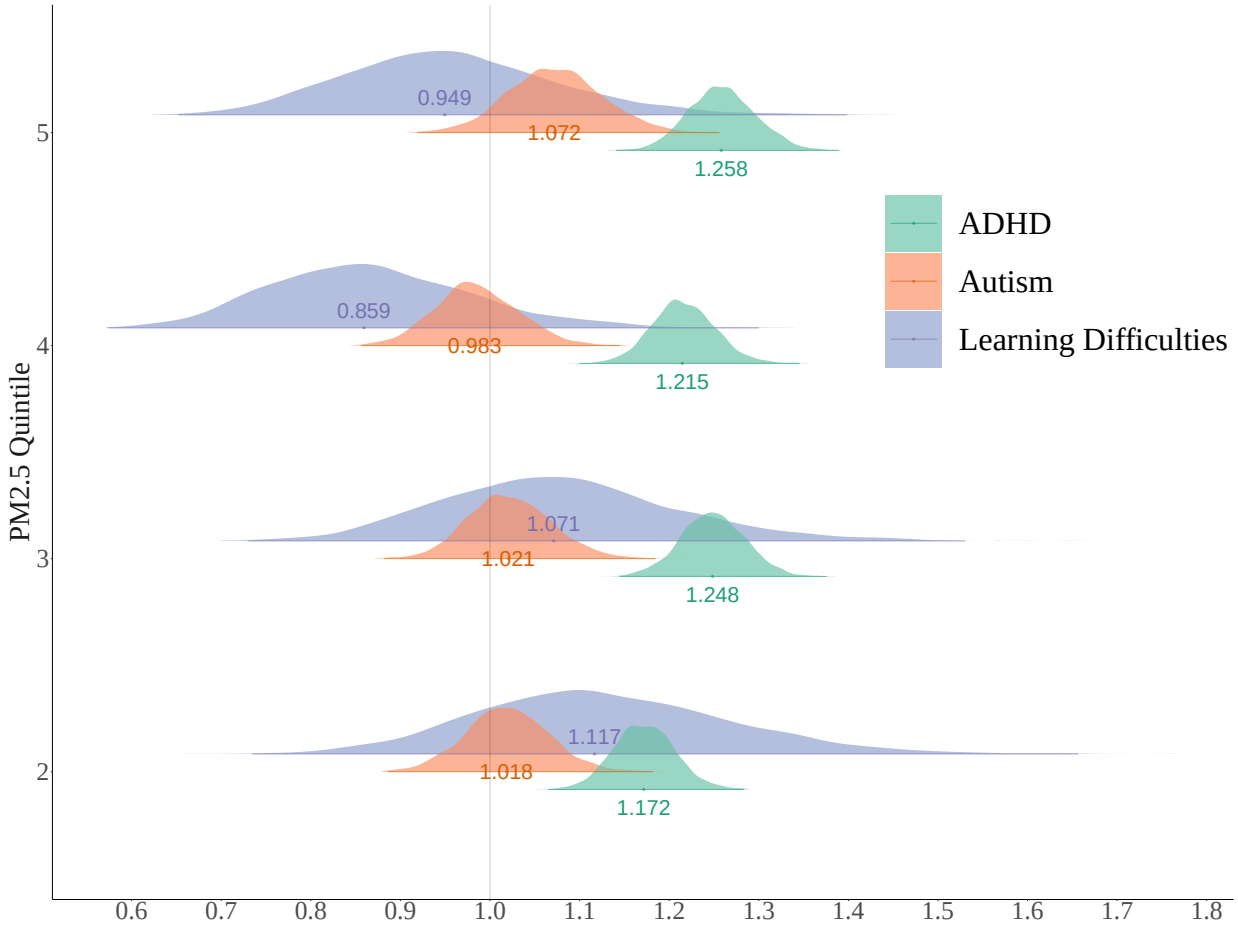


Figure 5.5: Posterior density estimation of OR $PM_{2.5}$ (M) Numerical annotations on the figure represent posterior medians, corresponding to the dots under the slabs. Both the slabs and the interval underneath represent 95% credible intervals.

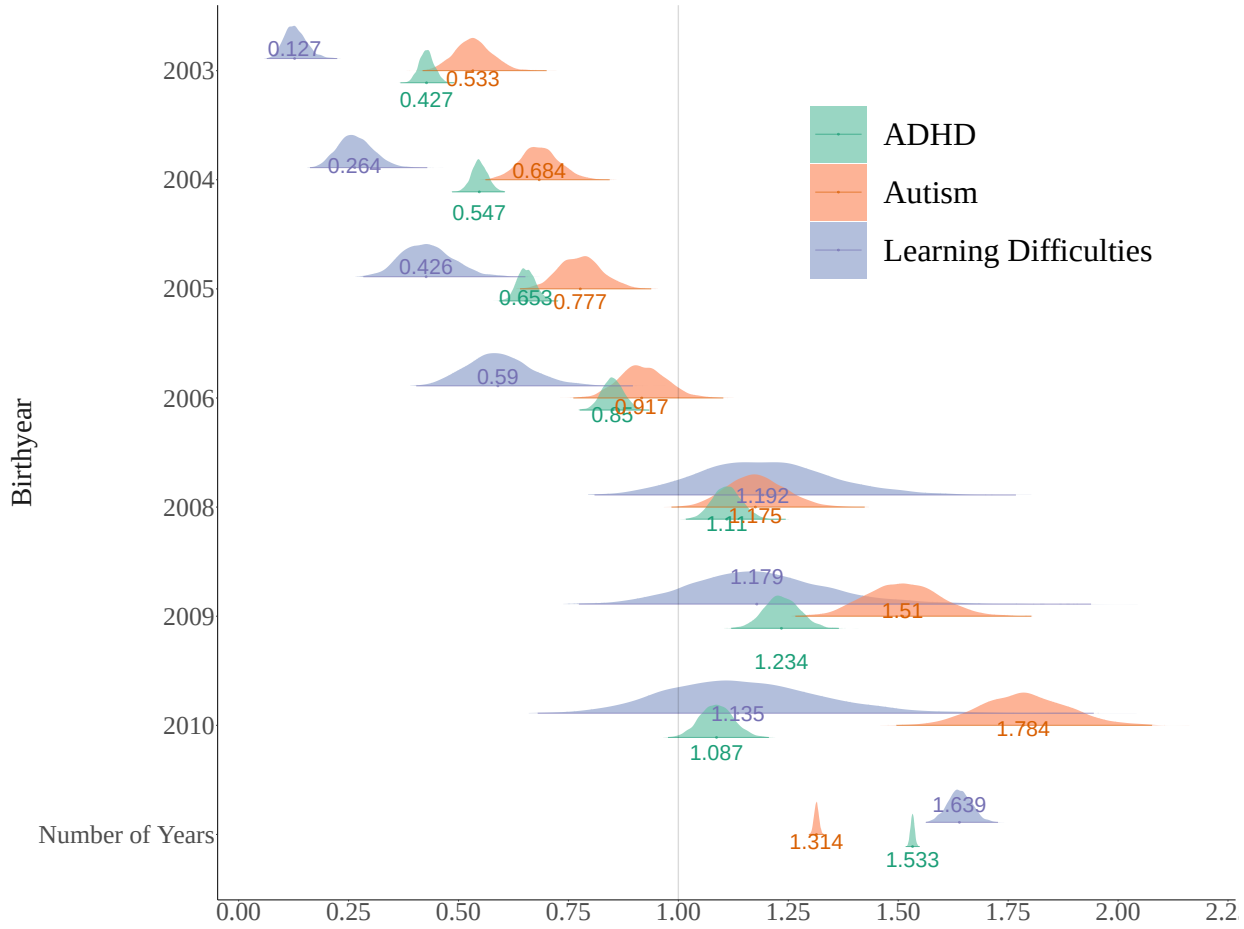


Figure 5.6: Posterior density estimation of OR for year of birth (B). The year 2007 was chosen as the reference level, and the level 2010 represent births both in 2010 and in 2011. Numerical annotations on the figure represent posterior medians, corresponding to the dots under the slabs. Both the slabs and the interval underneath represent 95% credible intervals.

Moreover, from 2004 to 2009 the ORs were monotonically increasing for all three outcomes, while for ASD this trend actually extended to 2010 as well. As one would expect, the longer a newborn was visible the more likely a diagnosis for any of the outcomes, as shown by elevated ORs for the offset term. Figure 5.7 show the mean of the posterior predictive distribution (PPD) on the logarithmic scale of the cubic P-splines smoothing terms over the day in year of birth (Wood, 2017). While the smooths for the learning difficulties model were not significant throughout the year, those for ADHD and ASD had peaks between August and September. Moreover, the ADHD smooths show another peak in spring as well.

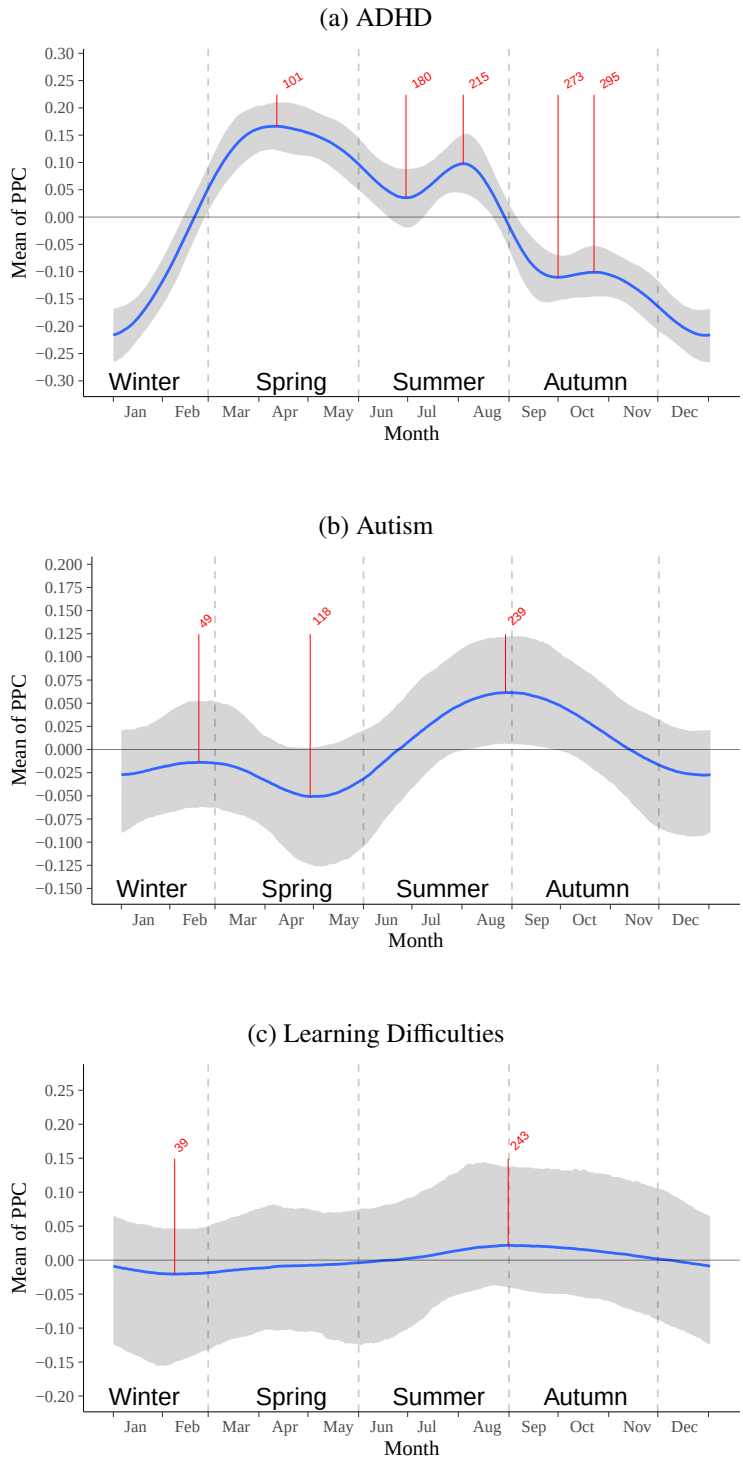


Figure 5.7: Mean of the posterior predictive distribution (PPD) and 95% credible intervals of the smoothing terms for the NDD submodels

5.4 Discussion

In this work we harnessed the power of a very large commercial health insurance claims dataset to probe the associations between maternal immune action, early-life immune system disorders, uses of anti-infective prescription, sex of newborn, abnormal birth conditions, time of birth in year and year of birth. The key advantage of our study is that we were able to account for the above risk factors within the same cohort and that our sample size was orders of magnitude larger than those from the existing literature. We now situate our results within the larger corpus of literature concerning putative aetiology of the NDDs in relation to the risk factors.

First, the importance of early life development of the microbiota has been increasingly recognized as central. For instance, (Olin et al., 2018) reported in a landmark study that although immune profiles (T cells and associated chemokines) of preterm and term newborns differed significantly at birth, with preterm newborns exhibiting elevated levels of inflammatory response, they eventually converged within 12 weeks, suggesting that it is early microbial interactions that drives immune system development, which can be hampered by dysbiosis. This is in turn supported by the observation that breast milk, to which preterm newborns generally have limited exposure, contains oligosaccharides that are digested by probiotic bacteria that stimulates the production of IgA (Henderickx et al., 2019). Among adults, on the other hand, it has been shown that the dynamics of immune cell reconstitution in cancer patients treated with allogeneic haematopoietic cell transplantation (HCT) was associated with the gut microbiome (Schluter et al., 2020).

Given the importance of the homeostasis of the microbiome, many studies have focused on the association between early life exposure to anti-infectives and NDDs. A general trend of decreased microbial diversity due to antibiotics use has been observed in children, with the resulting dysbiosis now known to be associated with a wide range of immune system-related diseases. Nevertheless, one should not hasten to draw a straightforward causal link between anti-infective use and NDDs since there may exist multiple confounding factors, especially maternal infections during pregnancy. For example, a discordant twins study using twin registers in the Netherlands and

Sweden showed that while early-life antibiotics use was associated with increased risks of ADHD and ASD, the effect size reduced considerably in the monozygotic (MZ) twins and same-sex dizygotic (DZ) sub-populations (Slob et al., 2021), despite twins being much more likely delivered via C-section than singletons — consistent with similar prospective twin studies on a Danish cohort (Axelsson et al., 2019a,b). This apparent inconsistency prompted a recent meta-analysis which confirmed the aforementioned results (Yu et al., 2022a). Nevertheless, since antibiotics could influence brain activity directly independent of dysbiosis, the question of the exact aetiology remains open (Champagne-Jorgensen et al., 2019).

Similarly, for C-section mode of delivery, the same twin study cited above (Axelsson et al., 2019a) showed that effect sizes of associations shrink towards zero when comparing siblings and twins, with a hazard ratio (HR) of 1.09 (95% CI 0.97–1.24) for intrapartum C-section, 1.03 (0.91–1.16) for prelabor C-section. Meanwhile, a meta-analysis of over 2.5 million subjects across 8 countries showed that the apparent association between C-sections (both elective and emergency) and ADHD disappeared when sibling data were pooled together (Xu et al., 2020). These results suggest that the causal factors were unmeasured maternal confounder(s) instead, which could include atopic diseases such as asthma, eczema or allergic rhinitis.

By contrast, the association between preterm births and ADHD and ASD yields a more straightforward pattern. A recent meta-analysis of 140 articles on early environmental risks showed that both preterm birth and low birth weight were associated with increased risks of NDDs even after familial confounding had been adjusted for (Carlsson et al., 2021). Moreover, case-control studies also pointed towards higher risks of ADHD: a Catalan cohort with 3,744 premature infants matched with 3,744 non-preterm had an HR of 5.52 (95% CI 1.73 – 17.62). The same pattern was also observed for the association between low birth weight and ADHD. In a population-based twin cohort with births in England and Wales between 1994 and 1996, the association between birth weight and ADHD symptomatology was statistically significant in an MZ-controlled setup, an effect persistent into adolescence and not moderated by gender, gestational age and low birth weight.

Finally, time of birth within a year was also associated with higher risks of ADHD. Consistent with previous literature, the relative risk of ADHD was found to be higher in spring and summer (more specifically, from March to August) and lower in autumn and winter (from September to February). The sharp drop in the curve between August and September has been commonly attributed to change in school year in many jurisdiction in the United States, since the youngest children in their cohort had the least opportunity for social interaction and physiological/psychological maturity, and hence higher probability of being diagnosed with NDDs (Root et al., 2019; Layton et al., 2018; Karlstad et al., 2017). What sets our result apart from the literature is the relative magnitude of the risk between March and August: we obtained a multimodal curve where the highest peak was located in March/April, whereas the literature generally puts the highest risk in late summer, around the beginning of school year (August/September) (Hsu et al., 2021). The last observation was, however, consistent with our findings for ASD, for which the risk was lowest in spring and highest in summer.

5.5 Appendix

5.5.1 Additional Methods

Regression Analysis

Immune System Submodels For these models, the outcome variable is the count of the given immune system-related disorder. Thus, the equation for the negative binomial (NB) regression is given by (Winkelmann, 2008, §§4.2.1 & 4.3, pp. 131 & 134)

$$\begin{aligned} y_i \sim \mathrm{NB}\left(r, \exp\left(\mathbf{x}^\top \boldsymbol{\beta}\right)\right) \iff \\ \mathbb{P}_{\mathrm{NB}}(y_i|\mathbf{x}, \boldsymbol{\beta}, r) = \frac{\Gamma(y_i + r)}{\Gamma(y_i - 1)\Gamma(r)} \left[\frac{\exp\left(\mathbf{x}^\top \boldsymbol{\beta}\right)}{\exp\left(\mathbf{x}^\top \boldsymbol{\beta}\right) + r}\right]^{y_i} \left[\frac{r}{\exp\left(\mathbf{x}^\top \boldsymbol{\beta}\right) + r}\right]^{r}. \end{aligned} \tag{5.1}$$

In count data regression, a regression coefficient (elements of $\boldsymbol{\beta}$) represents the relative effect on the expected counts of the outcome variable due to 1 unit increase of the corresponding predictor, or, in the case of binary predictors, the presence over the absence of the corresponding phenotype. Formally, the relative change in expected count, $\mathbb{E}(y|\cdot)$ (relative count change, ΔRC), due to such a change in $\mathbf{x}_j$ w.r.t. the baseline can be written as follows (Winkelmann, 2008, §3.1.4, pp. 70–71)

$$\begin{aligned} \Delta\mathrm{RC}_j &= \frac{\mathbb{E}\left(y|\,\mathbf{x}^{(j)\top}\boldsymbol{\beta}\right) - \mathbb{E}\left(y|\,\mathbf{x}^\top\boldsymbol{\beta}\right)}{\mathbb{E}\left(y|\,\mathbf{x}^\top\boldsymbol{\beta}\right)} \\ &= \frac{\exp\left(\mathbf{x}^\top\boldsymbol{\beta} + \boldsymbol{\beta}_j\right) - \exp\left(\mathbf{x}^\top\boldsymbol{\beta}\right)}{\exp\left(\mathbf{x}^\top\boldsymbol{\beta}\right)} \\ &= \exp\left(\boldsymbol{\beta}_j\right) - 1, \end{aligned} \tag{5.2}$$

where $\mathbf{x}^{(j)}$ is the same as $\mathbf{x}$ except that its j^{th} entry is increased by 1, and $\boldsymbol{\beta}_j$ the j^{th} entry of $\boldsymbol{\beta}_j$. In addition to the vanilla NB models we also fitted hurdle NB models, with likelihood given by

$$\mathbb{P}_{\text{Hurdle-NB}}(y_i|\mathbf{x}, \boldsymbol{\beta}, r) = \begin{cases} \pi & y_i = 0 \\ (1-\pi)\,\dfrac{\mathbb{P}_{\text{NB}}(y_i|\mathbf{x}, \boldsymbol{\beta}, r)}{1 - \mathbb{P}_{\text{NB}}(0|\mathbf{x}, \boldsymbol{\beta}, r)} & y_i > 0, \end{cases} \tag{5.3}$$

where $\pi \in [0, 1]$ represent the proportion of the data points $\{y_i\}$ equal to 0. As such, the hurdle model can be used to both inflated and deflated zero counts. For the prior on π we used $\text{Unif}(0, 1) \equiv \text{Beta}(1, 1)$.

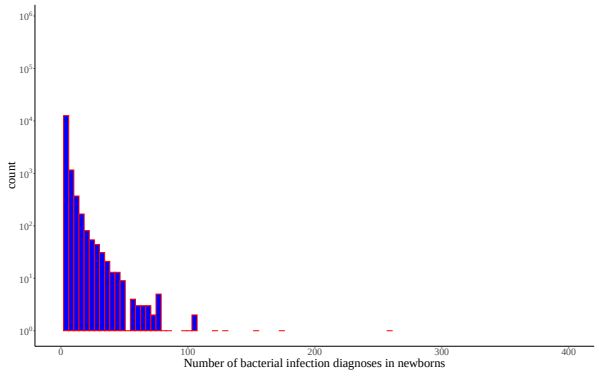
Over-dispersion Table 5.4 (see also Figure 5.8) shows that for each of the outcome phenotypes the variance was much larger than the mean, suggesting the presence of strong over-dispersion with respect to the Poisson distribution.

Phenotype	Median	Mean	Variance	Variance/Mean
Infections				
Bacterial	0	0.237	1.61	6.8
Viral	0	0.522	1.7	3.25
Miscellaneous	0	0.238	0.911	3.83
Sequelae	0	0.802	2.49	3.1
Immune disorder	2	5.49	181	32.9

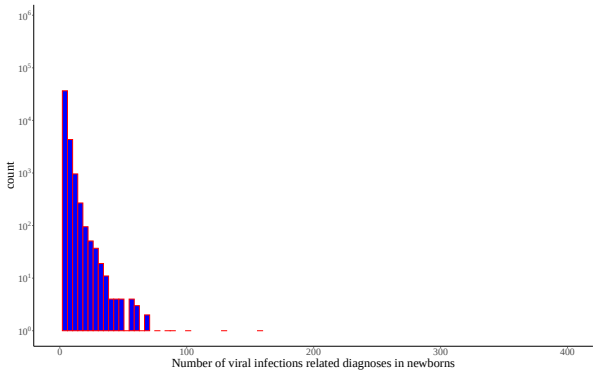
Table 5.4: Summary statistics of counts of the outcomes

Neurodevelopmental Submodels For these models, the outcome variable is the presence or absence of the given neurodevelopemntal disorder. We used logistic regression to model the probability of developing the given phenotype, such that the log-odds for the phenotype is the predicted by a linear combination of the predictors. The regression equation is given by

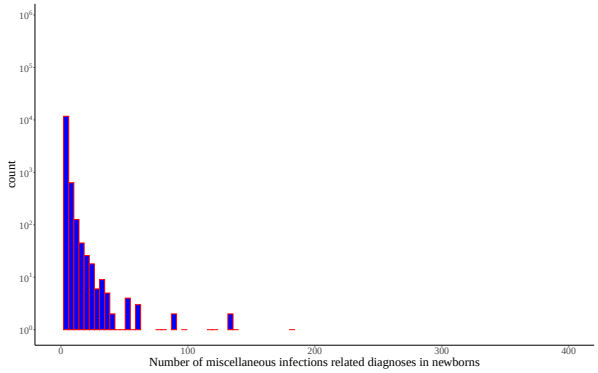
$$y_i = \text{logit}^{-1}\left(\mathbf{x}^\top \boldsymbol{\beta}\right), \qquad \text{where } \text{logit}^{-1}(z) = \frac{1}{1 + \exp(-z)} \tag{5.4}$$



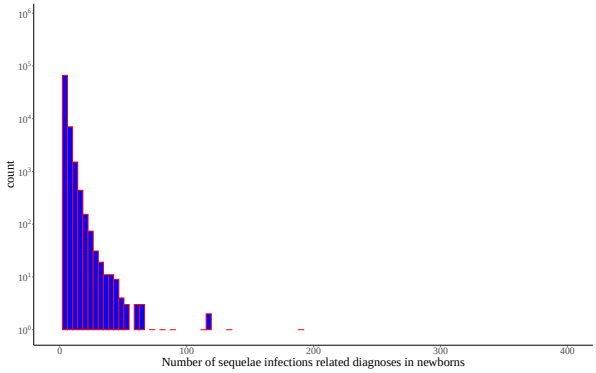
(a) Bacterial infections.



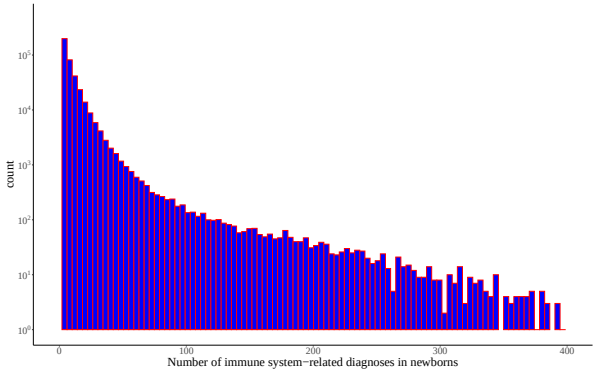
(b) Viral infections.



(c) Miscellaneous infections.



(d) Sequelae infections.



(e) Immune disorders.

Figure 5.8: Histograms of counts of phenotypes

In addition, to model the time of diagnosis we attempted to apply mixcure cure models from survival analysis. Such models consist of two parts: incidence and latency Amico and Van Keilegom (2018). The incidence part is identical to the logistic regression (Equation 5.4), and the latency part is modelled by a time-to-event model whose value is restricted to the positive real line. The posterior distributions from such models had weak identifiablity, as demonstrated by the multiple area of

high posterior density.

Smoothing Splines To model the effect of time of birth in year on the NDD outcomes, we used cyclic P-splines, which are defined as follows. In the first step, We define a set of B-splines serving as the basis for the smooths. Let x be the variable over which the smooths are placed, m be order of the splines (usually $m \geqslant 2$) and k be the number of spline parameters. Then we define $m+k+2$ knots $\{x_1, \cdots, x_{m+k+2}\}$, which are cutoff points on the range of the x's and are evenly spaced. The basis functions are then piecewise polynomials supported over the space spanned by $m + 2$ consecutive knots, given by the following recursive formula (base case is $m = -1$):

$$B_j^{(m)} = \frac{x - x_j}{x_{j+m+1} - x_j} B_j^{(m-1)}(x) + \frac{x_{j+m+2} - x}{x_{j+m+2} - x_{j+1}} B_j^{(m-1)}(x) \tag{5.5}$$

$$B_j^{(-1)} = \mathbb{1}(x)_{[x_j, x_{j+1})} \qquad j \in [1, k] \cap \mathbb{Z}. \tag{5.6}$$

As such, the m^{th} order B-splines are polynomial of order $m + 1$. The B-spline smooth can now be expressed as a linear combination of the basis functions

$$f(x) = \sum_{j=1}^{k} \beta_j B_j^{(m)}(x), \tag{5.7}$$

P-splines are essentially a regularized version of the B-splines, in that we incorporate additional terms that aim to penalize the L^2 distances between adjacent basis coefficients b_j

$$\mathcal{P} = \sum_{j=1}^{k-1} \left(\beta_{j+1} - \beta_j\right)^2 = \boldsymbol{\beta}^\top P^\top P \boldsymbol{\beta} = \boldsymbol{\beta}^\top S^\top \boldsymbol{\beta}, \tag{5.8}$$

where P has -1 on the diagonal, 1 on the superdiagonal and 0 everywhere else. Moreover, the smooth regression model

$$\mathbb{E}[y] = f(\mathbf{x}), \quad y \sim \mathcal{N}\left(0, \sigma^2\right) \tag{5.9}$$

can be recast in a hierarchical generalized linear model (HGLM) as

$$f(X) = X^\top \boldsymbol{\beta} + Z\mathbf{b}, \tag{5.10}$$

with $\mathbf{b} \sim$ i.i.d.$\mathcal{N}(0, \mathbf{I}\sigma_b^2)$, $X = QU^0$ and $Z = QU^+D^{-1/2}$, where we have performed the QR decomposition of $X = QR$, and $U = \begin{bmatrix} U^+ & U^0 \end{bmatrix}$ the matrix of eigenvectors of $R^{-\top}SR^{-1}$ with those corresponding to the positive eigenvalues U^+ and the rest U^0. Since the penalty matrix S is low rank, The diagonal matrix D contains the eigenvalues. There exist other forms of smoothing splines that do not necessitate choosing knots such as the thin plate regression splines (TPRS), which are more flexible and robust but unfortunately much more computationally expensive, rendering them impractical for our applications.

Statistical Modelling

Below are additional plots and tables for model comparisons.

model	elpd_diff	se_diff	elpd_kfold	se_elpd_kfold	p_kfold	se_p_kfold
C + P + W + A + R / H	0.000×10^{0}	0.000×10^{0}	-2.075×10^{5}	6.882×10^{2}	-1.636×10^{5}	4.290×10^{3}
C + P + A + R / H	-1.269×10^{2}	2.986×10^{1}	-2.076×10^{5}	6.882×10^{2}	-1.587×10^{5}	4.139×10^{3}
C + P + W + R / H	-1.835×10^{2}	3.691×10^{1}	-2.077×10^{5}	6.898×10^{2}	-1.636×10^{5}	4.299×10^{3}
C + P + W + A + R	-2.559×10^{2}	1.101×10^{2}	-2.077×10^{5}	7.117×10^{2}	-2.565×10^{5}	5.398×10^{3}
C + P + R / H	-3.422×10^{2}	4.725×10^{1}	-2.078×10^{5}	6.914×10^{2}	-1.586×10^{5}	4.137×10^{3}
C + P + A + R	-6.073×10^{2}	1.153×10^{2}	-2.081×10^{5}	7.138×10^{2}	-2.547×10^{5}	5.354×10^{3}
C + P + W + R	-7.461×10^{2}	1.234×10^{2}	-2.082×10^{5}	7.177×10^{2}	-2.568×10^{5}	5.431×10^{3}
C + P + A / H	-8.396×10^{2}	8.030×10^{1}	-2.083×10^{5}	6.949×10^{2}	-1.518×10^{5}	3.924×10^{3}
C + P + W / H	-8.992×10^{2}	8.675×10^{1}	-2.084×10^{5}	6.964×10^{2}	-1.560×10^{5}	4.057×10^{3}
C + P / H	-1.086×10^{3}	9.472×10^{1}	-2.086×10^{5}	6.994×10^{2}	-1.517×10^{5}	3.937×10^{3}
C + P + R	-1.113×10^{3}	1.300×10^{2}	-2.086×10^{5}	7.206×10^{2}	-2.549×10^{5}	5.374×10^{3}
C + P + W + A	-1.351×10^{3}	1.349×10^{2}	-2.088×10^{5}	7.193×10^{2}	-2.522×10^{5}	5.256×10^{3}
C + P + A	-1.697×10^{3}	1.407×10^{2}	-2.092×10^{5}	7.224×10^{2}	-2.505×10^{5}	5.211×10^{3}
C + P + W	-1.817×10^{3}	1.498×10^{2}	-2.093×10^{5}	7.265×10^{2}	-2.525×10^{5}	5.287×10^{3}
C + P	-2.160×10^{3}	1.559×10^{2}	-2.097×10^{5}	7.295×10^{2}	-2.508×10^{5}	5.241×10^{3}
C + P + W + A / H	–	–	–	–	–	–

Table 5.5: Model comparison for Immune System Submodels: Bacterial Infections

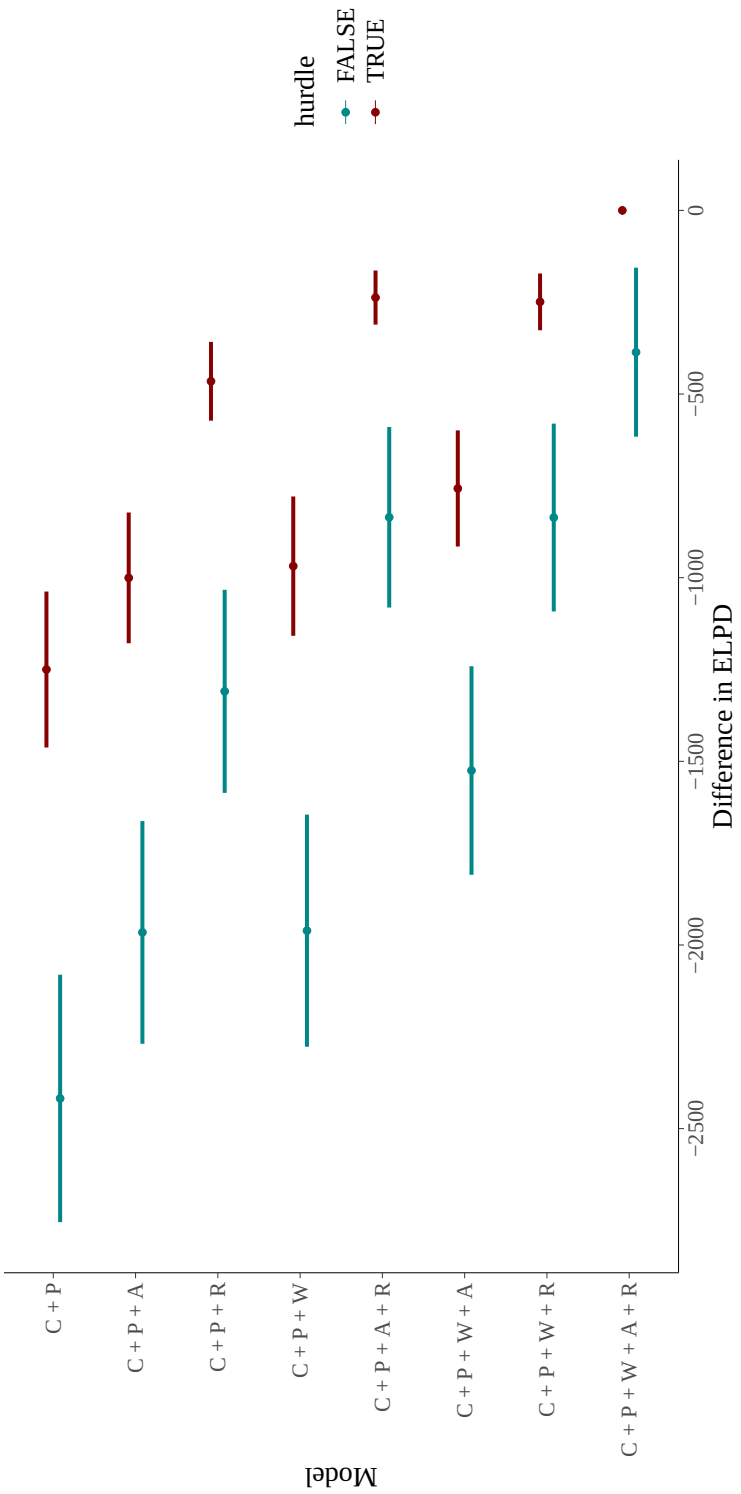


Figure 5.9: Model comparison for Immune System Submodels: Bacterial Infections

model	elpd_diff	se_diff	elpd_kfold	se_elpd_kfold	p_kfold	se_p_kfold
C + P + W + A + R / H	0.000×10^{0}	0.000×10^{0}	-7.860×10^{5}	9.034×10^{2}	-1.611×10^{6}	2.728×10^{4}
C + P + W + R / H	-5.428×10^{2}	5.082×10^{1}	-7.866×10^{5}	9.075×10^{2}	-1.617×10^{6}	2.754×10^{4}
C + P + A + R / H	-5.936×10^{2}	3.991×10^{1}	-7.866×10^{5}	9.068×10^{2}	-1.594×10^{6}	2.691×10^{4}
C + P + W + A / H	-1.063×10^{3}	6.613×10^{1}	-7.871×10^{5}	9.071×10^{2}	-1.520×10^{6}	2.475×10^{4}
C + P + R / H	-1.160×10^{3}	6.552×10^{1}	-7.872×10^{5}	9.124×10^{2}	-1.599×10^{6}	2.714×10^{4}
C + P + W / H	-1.625×10^{3}	8.400×10^{1}	-7.877×10^{5}	9.125×10^{2}	-1.526×10^{6}	2.503×10^{4}
C + P + A / H	-1.647×10^{3}	7.636×10^{1}	-7.877×10^{5}	9.105×10^{2}	-1.506×10^{6}	2.445×10^{4}
C + P / H	-2.222×10^{3}	9.315×10^{1}	-7.883×10^{5}	9.167×10^{2}	-1.512×10^{6}	2.470×10^{4}
C + P + W + A + R	-1.207×10^{4}	3.070×10^{2}	-7.981×10^{5}	1.055×10^{3}	-1.814×10^{6}	3.001×10^{4}
C + P + A + R	-1.309×10^{4}	3.249×10^{2}	-7.991×10^{5}	1.066×10^{3}	-1.799×10^{6}	2.968×10^{4}
C + P + W + R	-1.317×10^{4}	3.359×10^{2}	-7.992×10^{5}	1.071×10^{3}	-1.821×10^{6}	3.029×10^{4}
C + P + W + A	-1.397×10^{4}	3.334×10^{2}	-8.000×10^{5}	1.069×10^{3}	-1.720×10^{6}	2.727×10^{4}
C + P + R	-1.418×10^{4}	3.510×10^{2}	-8.002×10^{5}	1.082×10^{3}	-1.805×10^{6}	2.991×10^{4}
C + P + A	-1.501×10^{4}	3.528×10^{2}	-8.010×10^{5}	1.082×10^{3}	-1.709×10^{6}	2.702×10^{4}
C + P + W	-1.507×10^{4}	3.629×10^{2}	-8.011×10^{5}	1.087×10^{3}	-1.728×10^{6}	2.761×10^{4}
C + P	-1.608×10^{4}	3.796×10^{2}	-8.021×10^{5}	1.098×10^{3}	-1.716×10^{6}	2.737×10^{4}

Table 5.6: Model comparison for Immune System Submodels: Immune Disorders

5.5.2 Additional Results

Smoothing Splines

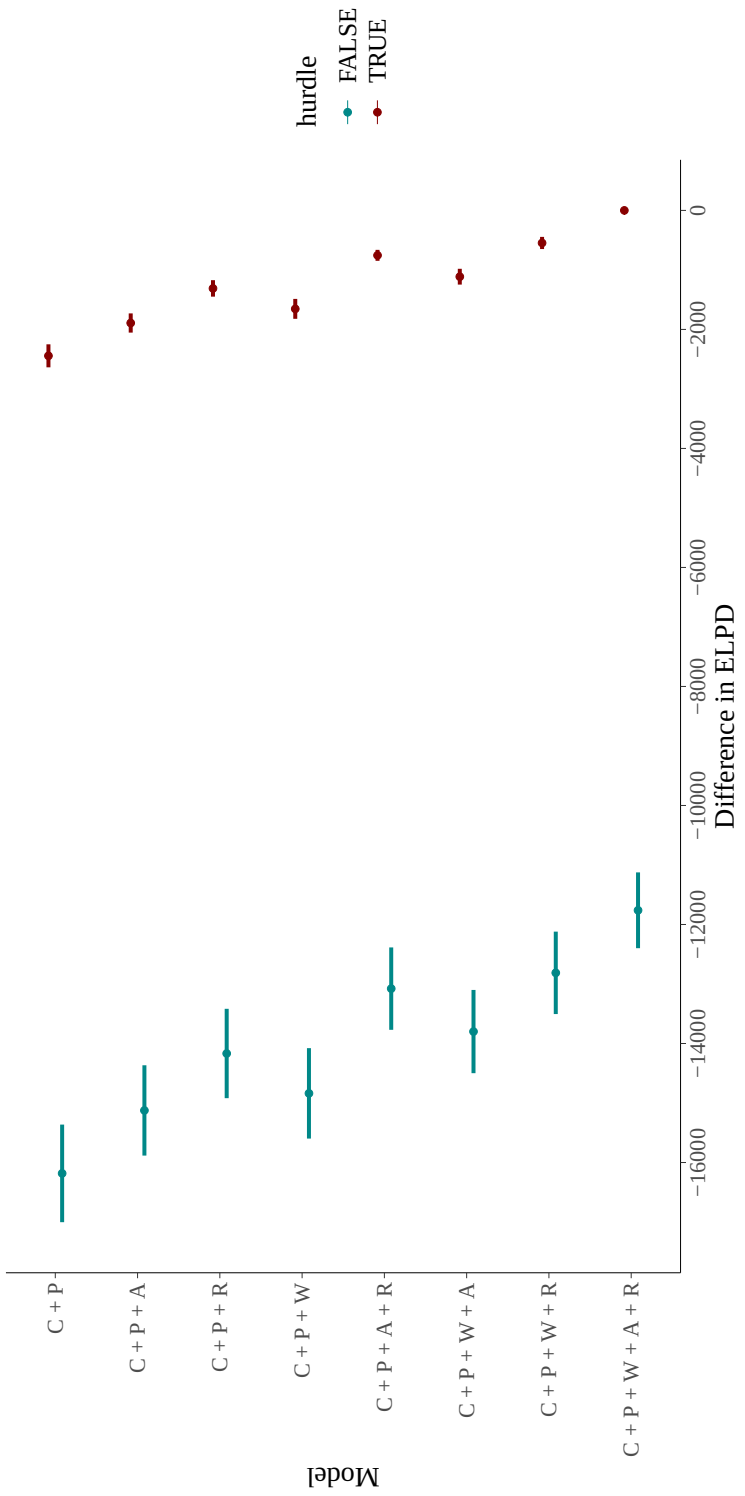


Figure 5.10: Model comparison for Immune System Submodels: Immune

model	elpd_diff	se_diff	elpd_kfold	se_elpd_kfold	p_kfold	se_p_kfold
C + P + W + A + R	0.000×10^{0}	0.000×10^{0}	-2.076×10^{5}	7.040×10^{2}	-2.141×10^{5}	4.086×10^{3}
C + P + W + A + R / H	-5.877×10^{1}	1.213×10^{2}	-2.077×10^{5}	6.735×10^{2}	-1.283×10^{5}	3.208×10^{3}
C + P + A + R / H	-2.122×10^{2}	1.252×10^{2}	-2.078×10^{5}	6.681×10^{2}	-1.271×10^{5}	3.181×10^{3}
C + P + W + R / H	-2.243×10^{2}	1.245×10^{2}	-2.079×10^{5}	6.712×10^{2}	-1.280×10^{5}	3.202×10^{3}
C + P + R / H	-4.038×10^{2}	1.315×10^{2}	-2.080×10^{5}	6.706×10^{2}	-1.269×10^{5}	3.176×10^{3}
C + P + W + R	-4.622×10^{2}	4.966×10^{1}	-2.081×10^{5}	7.046×10^{2}	-2.142×10^{5}	4.109×10^{3}
C + P + A + R	-4.808×10^{2}	4.440×10^{1}	-2.081×10^{5}	7.021×10^{2}	-2.135×10^{5}	4.080×10^{3}
C + P + W + A / H	-8.703×10^{2}	1.179×10^{2}	-2.085×10^{5}	6.768×10^{2}	-1.247×10^{5}	3.132×10^{3}
C + P + R	-9.625×10^{2}	7.418×10^{1}	-2.086×10^{5}	7.059×10^{2}	-2.135×10^{5}	4.098×10^{3}
C + P + W / H	-1.051×10^{3}	1.251×10^{2}	-2.087×10^{5}	6.809×10^{2}	-1.245×10^{5}	3.126×10^{3}
C + P + A / H	-1.074×10^{3}	1.284×10^{2}	-2.087×10^{5}	6.786×10^{2}	-1.237×10^{5}	3.109×10^{3}
C + P + W + A	-1.109×10^{3}	6.727×10^{1}	-2.087×10^{5}	7.141×10^{2}	-2.106×10^{5}	4.012×10^{3}
C + P / H	-1.272×10^{3}	1.357×10^{2}	-2.089×10^{5}	6.814×10^{2}	-1.234×10^{5}	3.104×10^{3}
C + P + W	-1.532×10^{3}	8.908×10^{1}	-2.092×10^{5}	7.171×10^{2}	-2.107×10^{5}	4.030×10^{3}
C + P + A	-1.552×10^{3}	8.691×10^{1}	-2.092×10^{5}	7.151×10^{2}	-2.101×10^{5}	4.006×10^{3}
C + P	-2.010×10^{3}	1.072×10^{2}	-2.096×10^{5}	7.197×10^{2}	-2.102×10^{5}	4.023×10^{3}

Table 5.7: Model comparison for Immune System Submodels: Miscellaneous Infections

model	elpd_diff	se_diff	elpd_kfold	se_elpd_kfold	p_kfold	se_p_kfold
C + P + W + A + R / H	0.000×10^{0}	0.000×10^{0}	-4.004×10^{5}	7.694×10^{2}	-7.192×10^{5}	1.556×10^{4}
C + P + W + R / H	-3.821×10^{2}	4.001×10^{1}	-4.007×10^{5}	7.761×10^{2}	-7.168×10^{5}	1.546×10^{4}
C + P + A + R / H	-5.583×10^{2}	3.773×10^{1}	-4.009×10^{5}	7.742×10^{2}	-7.120×10^{5}	1.536×10^{4}
C + P + R / H	-9.525×10^{2}	5.748×10^{1}	-4.013×10^{5}	7.813×10^{2}	-7.101×10^{5}	1.529×10^{4}
C + P + W + A / H	-2.013×10^{3}	8.841×10^{1}	-4.024×10^{5}	7.881×10^{2}	-7.064×10^{5}	1.524×10^{4}
C + P + W / H	-2.382×10^{3}	1.022×10^{2}	-4.027×10^{5}	7.951×10^{2}	-7.045×10^{5}	1.515×10^{4}
C + P + A / H	-2.609×10^{3}	1.014×10^{2}	-4.030×10^{5}	7.945×10^{2}	-6.999×10^{5}	1.506×10^{4}
C + P / H	-2.989×10^{3}	1.166×10^{2}	-4.033×10^{5}	8.030×10^{2}	-6.982×10^{5}	1.500×10^{4}
C + P + W + A + R	-8.212×10^{3}	2.386×10^{2}	-4.086×10^{5}	8.886×10^{2}	-8.049×10^{5}	1.559×10^{4}
C + P + W + R	-8.876×10^{3}	2.573×10^{2}	-4.092×10^{5}	9.001×10^{2}	-8.039×10^{5}	1.555×10^{4}
C + P + A + R	-9.088×10^{3}	2.495×10^{2}	-4.094×10^{5}	8.965×10^{2}	-8.000×10^{5}	1.547×10^{4}
C + P + R	-9.779×10^{3}	2.704×10^{2}	-4.101×10^{5}	9.096×10^{2}	-7.992×10^{5}	1.544×10^{4}
C + P + W + A	-1.113×10^{4}	2.874×10^{2}	-4.115×10^{5}	9.201×10^{2}	-7.945×10^{5}	1.533×10^{4}
C + P + W	-1.175×10^{4}	3.073×10^{2}	-4.121×10^{5}	9.332×10^{2}	-7.941×10^{5}	1.531×10^{4}
C + P + A	-1.199×10^{4}	3.003×10^{2}	-4.123×10^{5}	9.300×10^{2}	-7.903×10^{5}	1.523×10^{4}
C + P	-1.262×10^{4}	3.206×10^{2}	-4.130×10^{5}	9.436×10^{2}	-7.899×10^{5}	1.522×10^{4}

Table 5.8: Model comparison for Immune System Submodels: Sequelae Infections

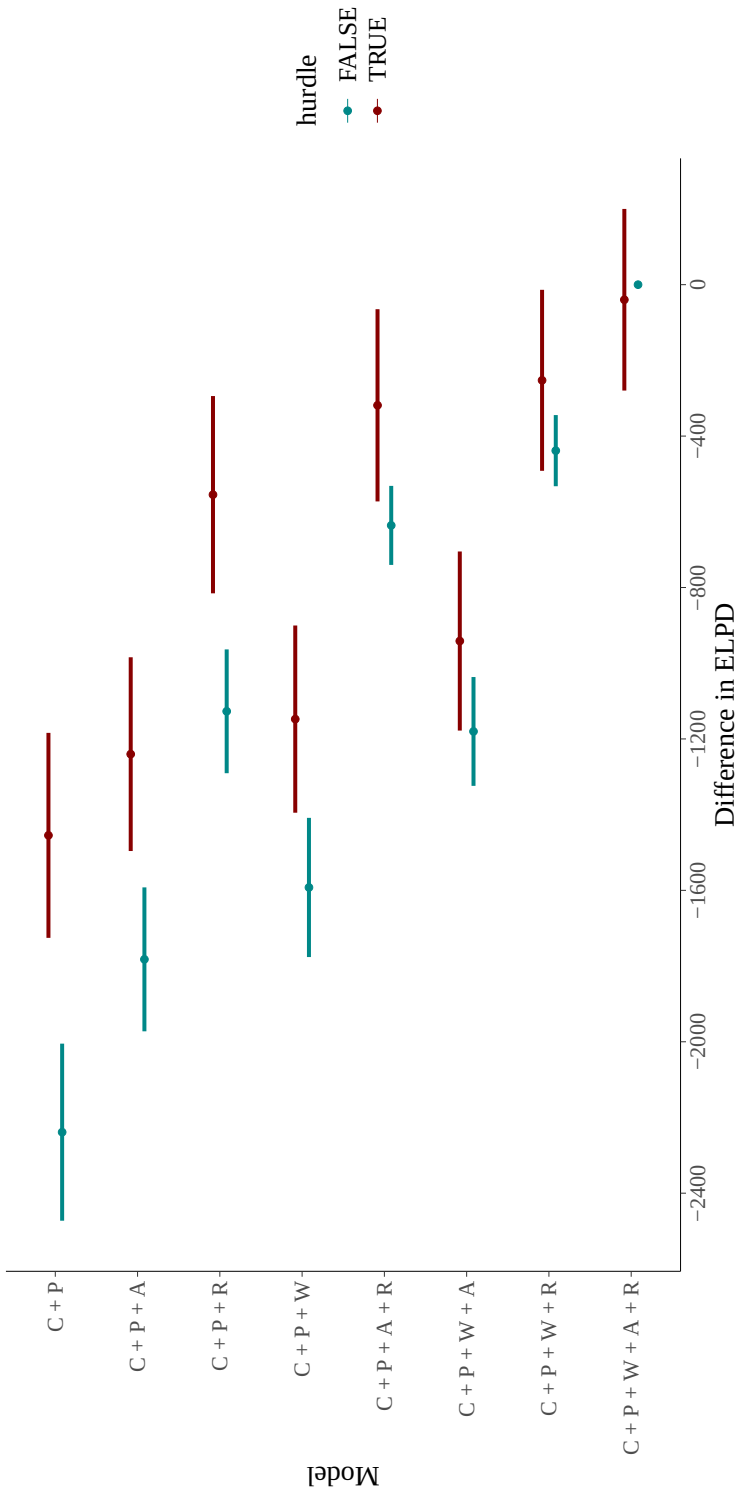


Figure 5.11: Model comparison for Immune System Submodels: Miscellaneous infections

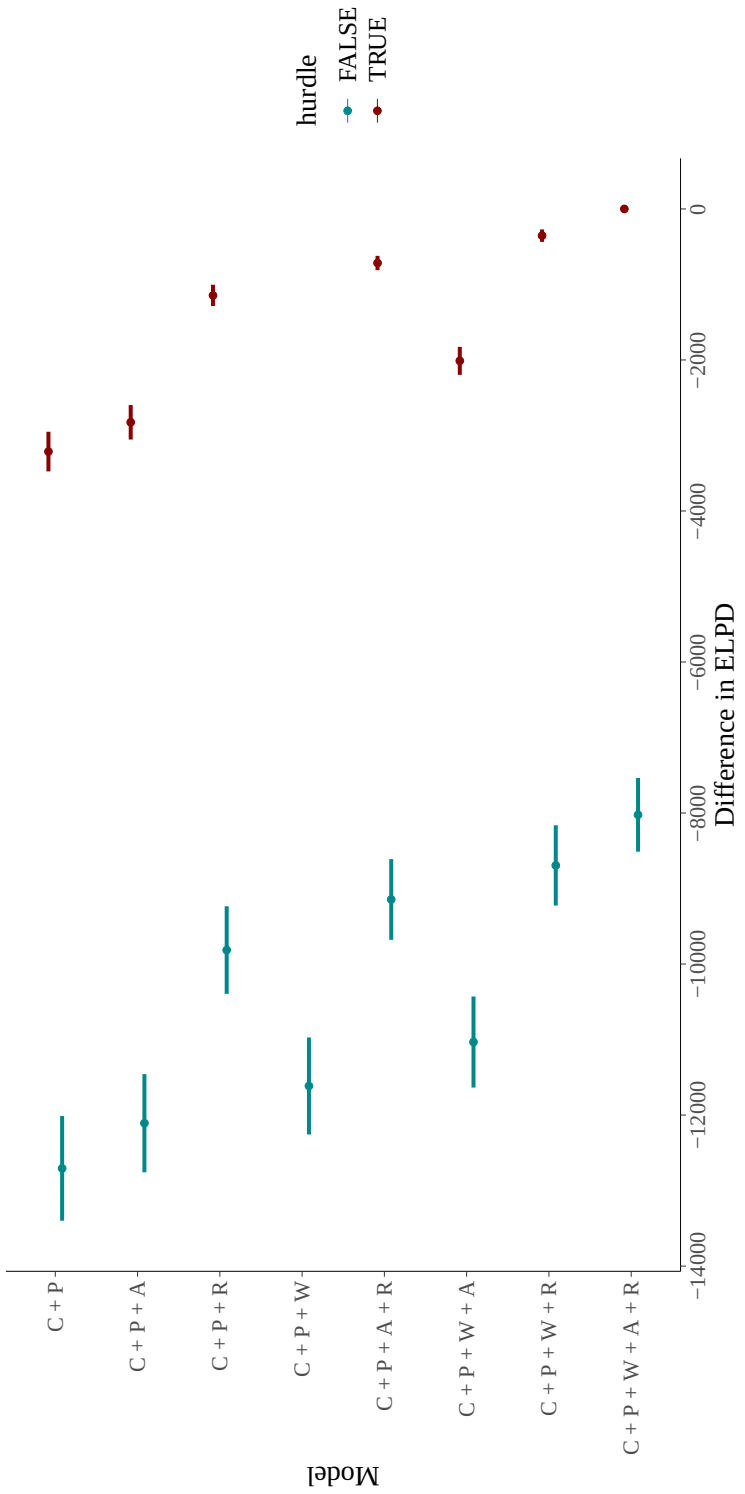


Figure 5.12: Model comparison for Immune System Submodels: Sequelae

model	elpd_diff	se_diff	elpd_kfold	se_elpd_kfold	p_kfold	se_p_kfold
C + P + W + A + R / H	0.000×10^{0}	0.000×10^{0}	-3.454×10^{5}	7.551×10^{2}	-5.436×10^{5}	1.121×10^{4}
C + P + W + R / H	-3.069×10^{2}	3.597×10^{1}	-3.457×10^{5}	7.593×10^{2}	-5.429×10^{5}	1.118×10^{4}
C + P + A + R / H	-5.622×10^{2}	4.583×10^{1}	-3.459×10^{5}	7.596×10^{2}	-5.417×10^{5}	1.118×10^{4}
C + P + R / H	-8.613×10^{2}	5.920×10^{1}	-3.462×10^{5}	7.644×10^{2}	-5.407×10^{5}	1.113×10^{4}
C + P + W + A / H	-1.193×10^{3}	6.973×10^{1}	-3.466×10^{5}	7.640×10^{2}	-5.283×10^{5}	1.076×10^{4}
C + P + W / H	-1.498×10^{3}	8.275×10^{1}	-3.469×10^{5}	7.700×10^{2}	-5.277×10^{5}	1.073×10^{4}
C + P + A / H	-1.768×10^{3}	8.489×10^{1}	-3.471×10^{5}	7.702×10^{2}	-5.266×10^{5}	1.073×10^{4}
C + P / H	-2.085×10^{3}	9.717×10^{1}	-3.474×10^{5}	7.751×10^{2}	-5.258×10^{5}	1.068×10^{4}
C + P + W + A + R	-4.159×10^{3}	1.756×10^{2}	-3.495×10^{5}	8.248×10^{2}	-6.394×10^{5}	1.159×10^{4}
C + P + W + R	-4.824×10^{3}	1.925×10^{2}	-3.502×10^{5}	8.342×10^{2}	-6.394×10^{5}	1.160×10^{4}
C + P + A + R	-5.090×10^{3}	1.875×10^{2}	-3.505×10^{5}	8.323×10^{2}	-6.381×10^{5}	1.157×10^{4}
C + P + R	-5.757×10^{3}	2.037×10^{2}	-3.511×10^{5}	8.415×10^{2}	-6.383×10^{5}	1.159×10^{4}
C + P + W + A	-5.943×10^{3}	2.046×10^{2}	-3.513×10^{5}	8.407×10^{2}	-6.277×10^{5}	1.122×10^{4}
C + P + W	-6.550×10^{3}	2.202×10^{2}	-3.519×10^{5}	8.498×10^{2}	-6.280×10^{5}	1.124×10^{4}
C + P + A	-6.828×10^{3}	2.161×10^{2}	-3.522×10^{5}	8.485×10^{2}	-6.269×10^{5}	1.121×10^{4}
C + P	-7.452×10^{3}	2.325×10^{2}	-3.528×10^{5}	8.584×10^{2}	-6.273×10^{5}	1.123×10^{4}

Table 5.9: Model comparison for Immune System Submodels: Viral Infections

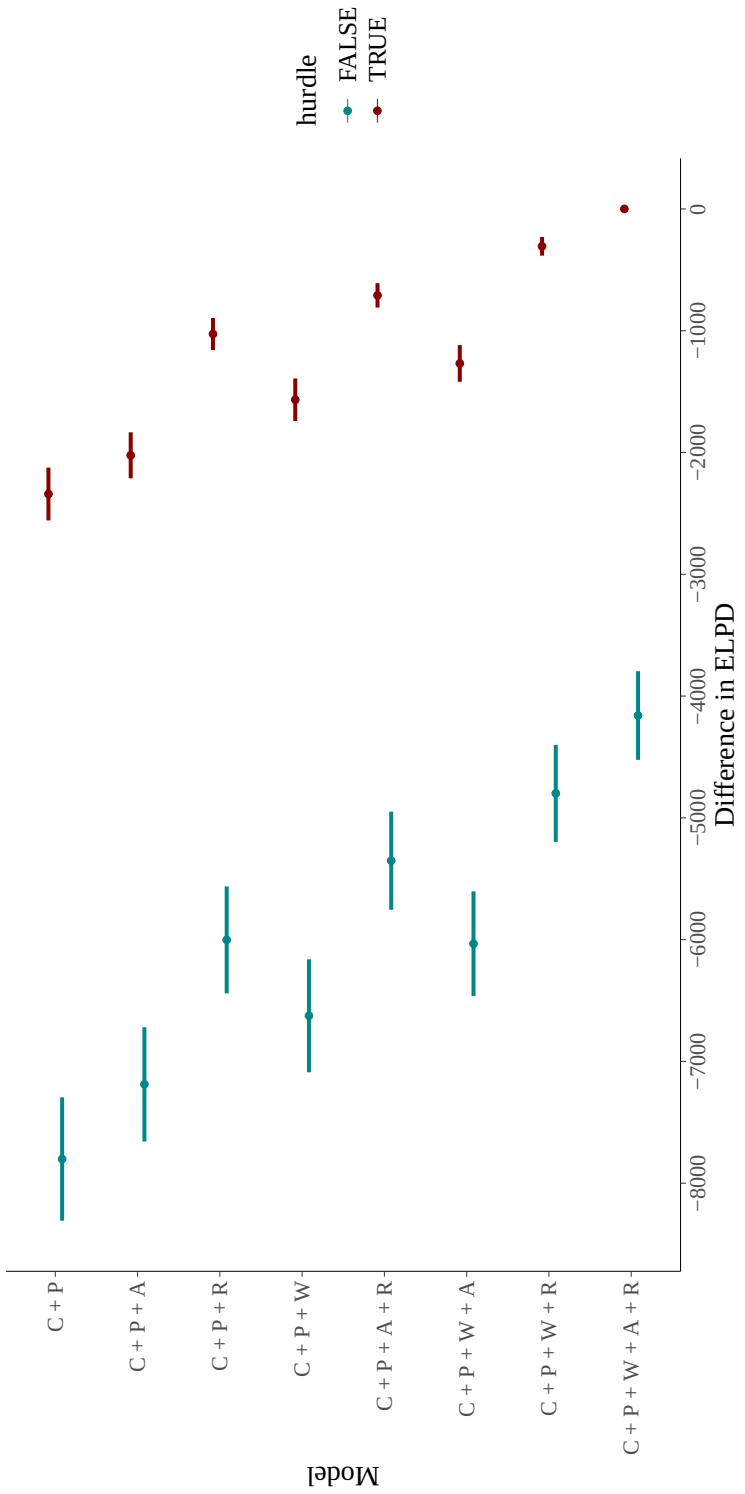


Figure 5.13: Model comparison for Immune System Submodels: Viral infection

model	elpd_diff	se_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
C + P + W + A + M + B + D + R + TD	0.000×10^{0}	0.000×10^{0}	-6.414×10^{4}	3.730×10^{2}	4.917×10^{1}	6.298×10^{-1}	1.283×10^{5}	7.460×10^{2}
C + P + W + A + M + B + D + R + TM	-5.111×10^{0}	4.844×10^{0}	-6.414×10^{4}	3.731×10^{2}	5.246×10^{1}	6.412×10^{-1}	1.283×10^{5}	7.462×10^{2}
C + P + W + M + B + D + R + TD	-2.240×10^{1}	7.021×10^{0}	-6.416×10^{4}	3.731×10^{2}	4.745×10^{1}	6.118×10^{-1}	1.283×10^{5}	7.462×10^{2}
C + P + A + M + B + D + R + TD	-2.502×10^{1}	7.858×10^{0}	-6.416×10^{4}	3.732×10^{2}	4.668×10^{1}	5.977×10^{-1}	1.283×10^{5}	7.463×10^{2}
C + P + W + M + B + D + R + TM	-2.796×10^{1}	8.555×10^{0}	-6.417×10^{4}	3.732×10^{2}	5.106×10^{1}	6.192×10^{-1}	1.283×10^{5}	7.464×10^{2}
C + P + A + M + B + D + R + TM	-3.051×10^{1}	9.207×10^{0}	-6.417×10^{4}	3.732×10^{2}	5.023×10^{1}	6.039×10^{-1}	1.283×10^{5}	7.464×10^{2}
C + P + M + B + D + R + TD	-4.748×10^{1}	1.047×10^{1}	-6.419×10^{4}	3.732×10^{2}	4.524×10^{1}	5.683×10^{-1}	1.284×10^{5}	7.465×10^{2}
C + P + M + B + D + R + TM	-5.242×10^{1}	1.154×10^{1}	-6.419×10^{4}	3.733×10^{2}	4.841×10^{1}	5.825×10^{-1}	1.284×10^{5}	7.466×10^{2}
C + P + W + A + M + B + D + R + TS	-6.163×10^{1}	1.032×10^{1}	-6.420×10^{4}	3.734×10^{2}	4.495×10^{1}	6.321×10^{-1}	1.284×10^{5}	7.467×10^{2}
C + P + W + M + B + D + R + TS	-8.317×10^{1}	1.254×10^{1}	-6.422×10^{4}	3.734×10^{2}	4.211×10^{1}	5.688×10^{-1}	1.284×10^{5}	7.469×10^{2}
C + P + A + M + B + D + R + TS	-8.618×10^{1}	1.291×10^{1}	-6.423×10^{4}	3.735×10^{2}	4.187×10^{1}	5.930×10^{-1}	1.285×10^{5}	7.470×10^{2}
C + P + M + B + D + R + TS	-1.082×10^{2}	1.470×10^{1}	-6.425×10^{4}	3.736×10^{2}	4.008×10^{1}	5.483×10^{-1}	1.285×10^{5}	7.471×10^{2}
C + P + W + A + M + B + D + TD	-1.247×10^{2}	1.663×10^{1}	-6.426×10^{4}	3.737×10^{2}	4.275×10^{1}	5.477×10^{-1}	1.285×10^{5}	7.473×10^{2}
C + P + W + A + M + B + D + TM	-1.294×10^{2}	1.730×10^{1}	-6.427×10^{4}	3.737×10^{2}	4.598×10^{1}	5.492×10^{-1}	1.285×10^{5}	7.474×10^{2}
C + P + W + M + B + D + TD	-1.481×10^{2}	1.804×10^{1}	-6.429×10^{4}	3.738×10^{2}	4.084×10^{1}	4.983×10^{-1}	1.286×10^{5}	7.475×10^{2}
C + P + A + M + B + D + TD	-1.512×10^{2}	1.838×10^{1}	-6.429×10^{4}	3.738×10^{2}	4.069×10^{1}	5.161×10^{-1}	1.286×10^{5}	7.476×10^{2}
C + P + W + M + B + D + TM	-1.527×10^{2}	1.868×10^{1}	-6.429×10^{4}	3.738×10^{2}	4.388×10^{1}	5.100×10^{-1}	1.286×10^{5}	7.477×10^{2}
C + P + A + M + B + D + TM	-1.562×10^{2}	1.898×10^{1}	-6.430×10^{4}	3.739×10^{2}	4.425×10^{1}	5.378×10^{-1}	1.286×10^{5}	7.477×10^{2}
C + P + M + B + D + TD	-1.741×10^{2}	1.963×10^{1}	-6.431×10^{4}	3.739×10^{2}	3.899×10^{1}	4.783×10^{-1}	1.286×10^{5}	7.478×10^{2}
C + P + M + B + D + TM	-1.787×10^{2}	2.021×10^{1}	-6.432×10^{4}	3.740×10^{2}	4.198×10^{1}	4.948×10^{-1}	1.286×10^{5}	7.479×10^{2}
C + P + W + A + M + B + D + TS	-1.832×10^{2}	1.960×10^{1}	-6.432×10^{4}	3.740×10^{2}	3.751×10^{1}	5.297×10^{-1}	1.286×10^{5}	7.480×10^{2}
C + P + W + M + B + D + TS	-2.069×10^{2}	2.084×10^{1}	-6.435×10^{4}	3.741×10^{2}	3.574×10^{1}	4.783×10^{-1}	1.287×10^{5}	7.481×10^{2}
C + P + A + M + B + D + TS	-2.099×10^{2}	2.107×10^{1}	-6.435×10^{4}	3.742×10^{2}	3.572×10^{1}	5.117×10^{-1}	1.287×10^{5}	7.483×10^{2}
C + P + W + A + M + B + R + TD	-2.175×10^{2}	2.211×10^{1}	-6.436×10^{4}	3.740×10^{2}	3.790×10^{1}	4.493×10^{-1}	1.287×10^{5}	7.480×10^{2}
C + P + W + A + M + B + R + TM	-2.223×10^{2}	2.268×10^{1}	-6.436×10^{4}	3.740×10^{2}	4.110×10^{1}	4.695×10^{-1}	1.287×10^{5}	7.480×10^{2}
C + P + M + B + D + TS	-2.328×10^{2}	2.220×10^{1}	-6.437×10^{4}	3.742×10^{2}	3.376×10^{1}	4.717×10^{-1}	1.287×10^{5}	7.485×10^{2}
C + P + W + M + B + R + TD	-2.406×10^{2}	2.327×10^{1}	-6.438×10^{4}	3.741×10^{2}	3.577×10^{1}	4.047×10^{-1}	1.288×10^{5}	7.482×10^{2}
C + P + A + M + B + R + TD	-2.447×10^{2}	2.355×10^{1}	-6.438×10^{4}	3.741×10^{2}	3.555×10^{1}	4.255×10^{-1}	1.288×10^{5}	7.482×10^{2}
C + P + W + M + B + R + TM	-2.453×10^{2}	2.381×10^{1}	-6.438×10^{4}	3.741×10^{2}	3.878×10^{1}	4.159×10^{-1}	1.288×10^{5}	7.482×10^{2}
C + P + A + M + B + R + TM	-2.497×10^{2}	2.407×10^{1}	-6.439×10^{4}	3.741×10^{2}	3.911×10^{1}	4.358×10^{-1}	1.288×10^{5}	7.483×10^{2}
C + P + M + B + R + TD	-2.675×10^{2}	2.462×10^{1}	-6.441×10^{4}	3.742×10^{2}	3.370×10^{1}	3.702×10^{-1}	1.288×10^{5}	7.484×10^{2}
C + P + M + B + R + TM	-2.723×10^{2}	2.513×10^{1}	-6.441×10^{4}	3.742×10^{2}	3.699×10^{1}	3.791×10^{-1}	1.288×10^{5}	7.485×10^{2}
C + P + W + A + M + B + R + TS	-2.770×10^{2}	2.427×10^{1}	-6.442×10^{4}	3.743×10^{2}	3.249×10^{1}	4.426×10^{-1}	1.288×10^{5}	7.485×10^{2}
C + P + W + M + B + R + TS	-3.009×10^{2}	2.536×10^{1}	-6.444×10^{4}	3.744×10^{2}	3.105×10^{1}	3.852×10^{-1}	1.289×10^{5}	7.487×10^{2}
C + P + A + M + B + R + TS	-3.046×10^{2}	2.556×10^{1}	-6.444×10^{4}	3.744×10^{2}	3.070×10^{1}	4.069×10^{-1}	1.289×10^{5}	7.488×10^{2}
C + P + M + B + R + TS	-3.270×10^{2}	2.658×10^{1}	-6.447×10^{4}	3.745×10^{2}	2.832×10^{1}	3.535×10^{-1}	1.289×10^{5}	7.490×10^{2}
C + P + W + A + M + B + TD	-4.970×10^{2}	3.315×10^{1}	-6.464×10^{4}	3.753×10^{2}	3.180×10^{1}	3.326×10^{-1}	1.293×10^{5}	7.507×10^{2}
C + P + W + A + M + B + TM	-5.007×10^{2}	3.353×10^{1}	-6.464×10^{4}	3.754×10^{2}	3.444×10^{1}	3.462×10^{-1}	1.293×10^{5}	7.508×10^{2}
C + P + W + M + B + TD	-5.202×10^{2}	3.389×10^{1}	-6.466×10^{4}	3.754×10^{2}	2.947×10^{1}	2.693×10^{-1}	1.293×10^{5}	7.509×10^{2}
C + P + W + M + B + TM	-5.247×10^{2}	3.427×10^{1}	-6.466×10^{4}	3.755×10^{2}	3.280×10^{1}	2.876×10^{-1}	1.293×10^{5}	7.510×10^{2}
C + P + A + M + B + TD	-5.262×10^{2}	3.413×10^{1}	-6.467×10^{4}	3.755×10^{2}	2.948×10^{1}	2.983×10^{-1}	1.293×10^{5}	7.509×10^{2}
C + P + A + M + B + TM	-5.304×10^{2}	3.450×10^{1}	-6.467×10^{4}	3.755×10^{2}	3.263×10^{1}	3.105×10^{-1}	1.293×10^{5}	7.511×10^{2}
C + P + M + B + TD	-5.487×10^{2}	3.484×10^{1}	-6.469×10^{4}	3.756×10^{2}	2.710×10^{1}	2.143×10^{-1}	1.294×10^{5}	7.511×10^{2}
C + P + M + B + TM	-5.534×10^{2}	3.521×10^{1}	-6.469×10^{4}	3.756×10^{2}	3.051×10^{1}	2.361×10^{-1}	1.294×10^{5}	7.513×10^{2}
C + P + W + A + M + B + TS	-5.539×10^{2}	3.464×10^{1}	-6.469×10^{4}	3.757×10^{2}	2.614×10^{1}	3.099×10^{-1}	1.294×10^{5}	7.513×10^{2}
C + P + W + M + B + TS	-5.780×10^{2}	3.537×10^{1}	-6.472×10^{4}	3.757×10^{2}	2.464×10^{1}	2.457×10^{-1}	1.294×10^{5}	7.515×10^{2}
C + P + A + M + B + TS	-5.836×10^{2}	3.557×10^{1}	-6.472×10^{4}	3.758×10^{2}	2.444×10^{1}	2.752×10^{-1}	1.294×10^{5}	7.517×10^{2}
C + P + M + B + TS	-6.067×10^{2}	3.626×10^{1}	-6.475×10^{4}	3.759×10^{2}	2.239×10^{1}	1.885×10^{-1}	1.295×10^{5}	7.518×10^{2}

Table 5.10: Model comparison for NDD Submodels:ADHD

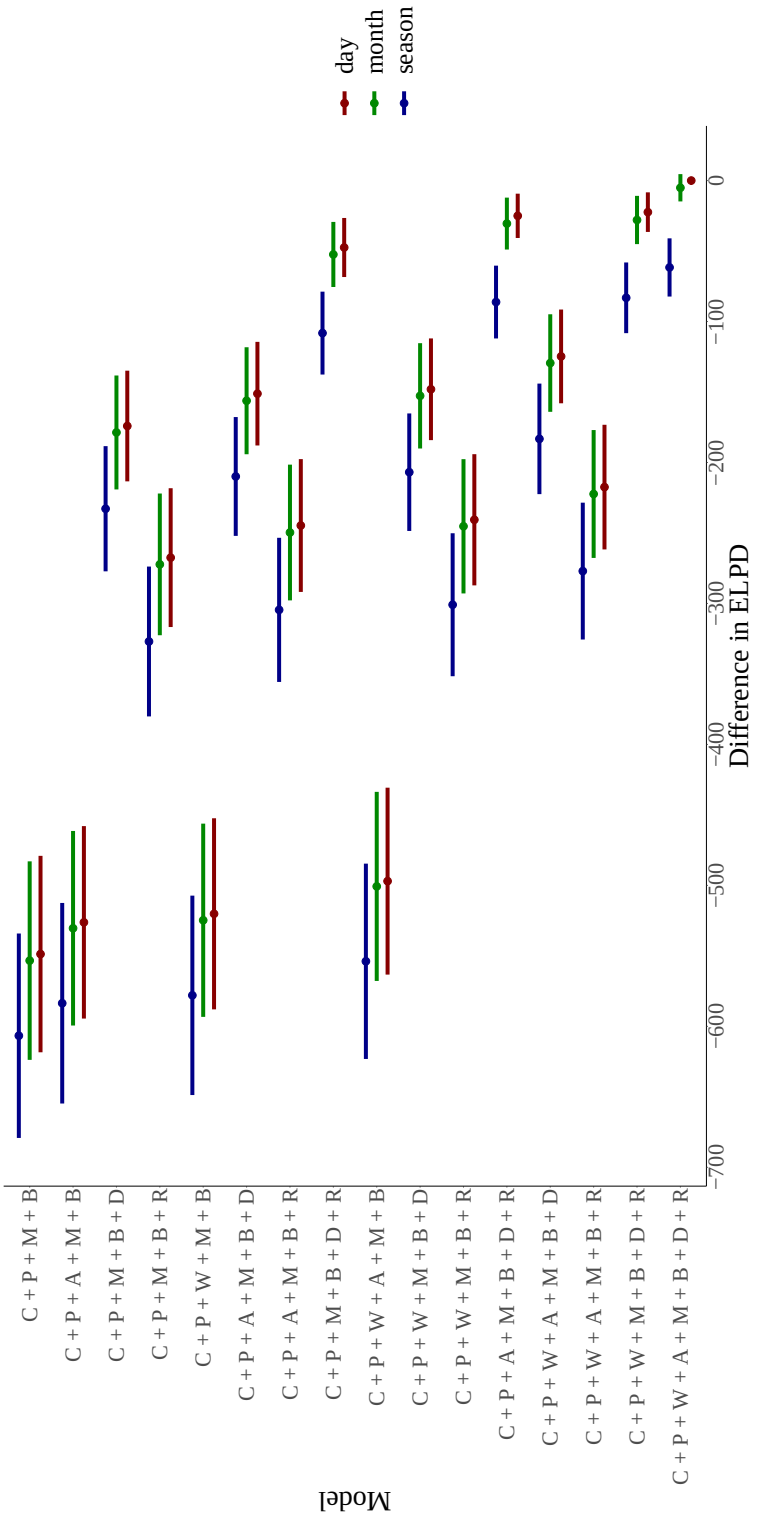


Figure 5.14: Model comparison for Neurodvelopmental submodels: ADHD

model	elpd_diff	se_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
C + P + W + A + M + B + D + R + TS	0.000×10^{0}	0.000×10^{0}	-2.764×10^{4}	3.211×10^{2}	4.283×10^{1}	1.103×10^{0}	5.528×10^{4}	6.422×10^{2}
C + P + W + A + M + B + D + R + TD	-1.267×10^{-1}	1.892×10^{0}	-2.764×10^{4}	3.211×10^{2}	4.435×10^{1}	1.126×10^{0}	5.529×10^{4}	6.421×10^{2}
C + P + W + A + M + B + D + TS	-2.984×10^{0}	4.152×10^{0}	-2.765×10^{4}	3.211×10^{2}	3.684×10^{1}	9.321×10^{-1}	5.529×10^{4}	6.422×10^{2}
C + P + W + A + M + B + D + TD	-3.283×10^{0}	4.584×10^{0}	-2.765×10^{4}	3.211×10^{2}	3.821×10^{1}	9.571×10^{-1}	5.529×10^{4}	6.422×10^{2}
C + P + W + A + M + B + D + R + TM	-4.678×10^{0}	2.968×10^{0}	-2.765×10^{4}	3.212×10^{2}	5.184×10^{1}	1.205×10^{0}	5.529×10^{4}	6.423×10^{2}
C + P + W + A + M + B + D + TM	-7.369×10^{0}	5.106×10^{0}	-2.765×10^{4}	3.212×10^{2}	4.548×10^{1}	1.033×10^{0}	5.530×10^{4}	6.425×10^{2}
C + P + W + A + M + B + R + TS	-1.848×10^{1}	7.839×10^{0}	-2.766×10^{4}	3.213×10^{2}	3.185×10^{1}	8.119×10^{-1}	5.532×10^{4}	6.426×10^{2}
C + P + W + A + M + B + R + TD	-1.871×10^{1}	8.056×10^{0}	-2.766×10^{4}	3.213×10^{2}	3.358×10^{1}	8.361×10^{-1}	5.532×10^{4}	6.425×10^{2}
C + P + W + A + M + B + TD	-1.937×10^{1}	8.867×10^{0}	-2.766×10^{4}	3.213×10^{2}	2.720×10^{1}	6.009×10^{-1}	5.532×10^{4}	6.426×10^{2}
C + P + W + A + M + B + TS	-1.951×10^{1}	8.664×10^{0}	-2.766×10^{4}	3.213×10^{2}	2.614×10^{1}	6.014×10^{-1}	5.532×10^{4}	6.426×10^{2}
C + P + W + A + M + B + R + TM	-2.265×10^{1}	8.400×10^{0}	-2.767×10^{4}	3.213×10^{2}	4.034×10^{1}	8.658×10^{-1}	5.533×10^{4}	6.427×10^{2}
C + P + W + A + M + B + TM	-2.350×10^{1}	9.180×10^{0}	-2.767×10^{4}	3.213×10^{2}	3.435×10^{1}	6.732×10^{-1}	5.533×10^{4}	6.427×10^{2}
C + P + A + M + B + D + R + TD	-3.116×10^{1}	8.876×10^{0}	-2.767×10^{4}	3.214×10^{2}	4.239×10^{1}	1.090×10^{0}	5.535×10^{4}	6.427×10^{2}
C + P + A + M + B + D + R + TS	-3.122×10^{1}	8.686×10^{0}	-2.767×10^{4}	3.214×10^{2}	4.121×10^{1}	1.085×10^{0}	5.535×10^{4}	6.428×10^{2}
C + P + A + M + B + D + TD	-3.419×10^{1}	9.836×10^{0}	-2.768×10^{4}	3.214×10^{2}	3.612×10^{1}	9.285×10^{-1}	5.535×10^{4}	6.428×10^{2}
C + P + A + M + B + D + TS	-3.420×10^{1}	9.654×10^{0}	-2.768×10^{4}	3.214×10^{2}	3.513×10^{1}	9.140×10^{-1}	5.535×10^{4}	6.429×10^{2}
C + P + W + M + B + D + R + TD	-3.455×10^{1}	8.887×10^{0}	-2.768×10^{4}	3.215×10^{2}	4.199×10^{1}	1.026×10^{0}	5.535×10^{4}	6.430×10^{2}
C + P + A + M + B + D + R + TM	-3.468×10^{1}	9.154×10^{0}	-2.768×10^{4}	3.215×10^{2}	4.902×10^{1}	1.149×10^{0}	5.535×10^{4}	6.429×10^{2}
C + P + W + M + B + D + R + TS	-3.504×10^{1}	8.716×10^{0}	-2.768×10^{4}	3.215×10^{2}	4.112×10^{1}	1.041×10^{0}	5.535×10^{4}	6.430×10^{2}
C + P + A + M + B + D + TM	-3.770×10^{1}	1.008×10^{1}	-2.768×10^{4}	3.215×10^{2}	4.282×10^{1}	9.796×10^{-1}	5.536×10^{4}	6.430×10^{2}
C + P + W + M + B + D + TS	-3.830×10^{1}	9.745×10^{0}	-2.768×10^{4}	3.215×10^{2}	3.469×10^{1}	8.216×10^{-1}	5.536×10^{4}	6.431×10^{2}
C + P + W + M + B + D + R + TM	-3.857×10^{1}	9.206×10^{0}	-2.768×10^{4}	3.215×10^{2}	4.906×10^{1}	1.095×10^{0}	5.536×10^{4}	6.431×10^{2}
C + P + W + M + B + D + TD	-3.868×10^{1}	9.910×10^{0}	-2.768×10^{4}	3.215×10^{2}	3.614×10^{1}	8.610×10^{-1}	5.536×10^{4}	6.430×10^{2}
C + P + W + M + B + D + TM	-4.226×10^{1}	1.019×10^{1}	-2.768×10^{4}	3.216×10^{2}	4.297×10^{1}	9.063×10^{-1}	5.537×10^{4}	6.432×10^{2}
C + P + A + M + B + R + TS	-4.984×10^{1}	1.170×10^{1}	-2.769×10^{4}	3.215×10^{2}	2.966×10^{1}	7.504×10^{-1}	5.538×10^{4}	6.431×10^{2}
C + P + A + M + B + R + TD	-5.032×10^{1}	1.183×10^{1}	-2.769×10^{4}	3.215×10^{2}	3.141×10^{1}	7.992×10^{-1}	5.539×10^{4}	6.431×10^{2}
C + P + A + M + B + TS	-5.102×10^{1}	1.231×10^{1}	-2.769×10^{4}	3.216×10^{2}	2.390×10^{1}	5.576×10^{-1}	5.539×10^{4}	6.431×10^{2}
C + P + A + M + B + TD	-5.127×10^{1}	1.245×10^{1}	-2.769×10^{4}	3.216×10^{2}	2.535×10^{1}	5.680×10^{-1}	5.539×10^{4}	6.432×10^{2}
C + P + W + M + B + R + TD	-5.271×10^{1}	1.184×10^{1}	-2.770×10^{4}	3.216×10^{2}	3.138×10^{1}	7.212×10^{-1}	5.539×10^{4}	6.433×10^{2}
C + P + W + M + B + R + TS	-5.295×10^{1}	1.172×10^{1}	-2.770×10^{4}	3.217×10^{2}	3.010×10^{1}	7.005×10^{-1}	5.539×10^{4}	6.433×10^{2}
C + P + W + M + B + TD	-5.390×10^{1}	1.246×10^{1}	-2.770×10^{4}	3.217×10^{2}	2.511×10^{1}	4.442×10^{-1}	5.539×10^{4}	6.433×10^{2}
C + P + A + M + B + R + TM	-5.404×10^{1}	1.207×10^{1}	-2.770×10^{4}	3.217×10^{2}	3.814×10^{1}	8.556×10^{-1}	5.539×10^{4}	6.433×10^{2}
C + P + W + M + B + TS	-5.446×10^{1}	1.234×10^{1}	-2.770×10^{4}	3.217×10^{2}	2.425×10^{1}	4.422×10^{-1}	5.539×10^{4}	6.434×10^{2}
C + P + A + M + B + TM	-5.492×10^{1}	1.265×10^{1}	-2.770×10^{4}	3.216×10^{2}	3.198×10^{1}	6.257×10^{-1}	5.539×10^{4}	6.433×10^{2}
C + P + W + M + B + R + TM	-5.643×10^{1}	1.211×10^{1}	-2.770×10^{4}	3.217×10^{2}	3.800×10^{1}	7.609×10^{-1}	5.540×10^{4}	6.434×10^{2}
C + P + W + M + B + TM	-5.780×10^{1}	1.270×10^{1}	-2.770×10^{4}	3.217×10^{2}	3.189×10^{1}	5.245×10^{-1}	5.540×10^{4}	6.435×10^{2}
C + P + M + B + D + R + TS	-6.671×10^{1}	1.253×10^{1}	-2.771×10^{4}	3.218×10^{2}	3.887×10^{1}	1.001×10^{0}	5.542×10^{4}	6.436×10^{2}
C + P + M + B + D + R + TD	-6.679×10^{1}	1.264×10^{1}	-2.771×10^{4}	3.218×10^{2}	4.019×10^{1}	1.019×10^{0}	5.542×10^{4}	6.435×10^{2}
C + P + M + B + D + R + TM	-7.045×10^{1}	1.287×10^{1}	-2.771×10^{4}	3.219×10^{2}	4.700×10^{1}	1.079×10^{0}	5.543×10^{4}	6.437×10^{2}
C + P + M + B + D + TD	-7.063×10^{1}	1.340×10^{1}	-2.771×10^{4}	3.218×10^{2}	3.388×10^{1}	7.951×10^{-1}	5.543×10^{4}	6.437×10^{2}
C + P + M + B + D + TS	-7.083×10^{1}	1.329×10^{1}	-2.771×10^{4}	3.218×10^{2}	3.316×10^{1}	8.033×10^{-1}	5.543×10^{4}	6.437×10^{2}
C + P + M + B + D + TM	-7.439×10^{1}	1.360×10^{1}	-2.772×10^{4}	3.219×10^{2}	4.101×10^{1}	8.758×10^{-1}	5.543×10^{4}	6.438×10^{2}
C + P + M + B + R + TS	-8.530×10^{1}	1.478×10^{1}	-2.773×10^{4}	3.220×10^{2}	2.781×10^{1}	6.592×10^{-1}	5.546×10^{4}	6.440×10^{2}
C + P + M + B + R + TD	-8.540×10^{1}	1.487×10^{1}	-2.773×10^{4}	3.220×10^{2}	2.934×10^{1}	6.894×10^{-1}	5.546×10^{4}	6.439×10^{2}
C + P + M + B + TS	-8.701×10^{1}	1.532×10^{1}	-2.773×10^{4}	3.220×10^{2}	2.193×10^{1}	3.680×10^{-1}	5.546×10^{4}	6.440×10^{2}
C + P + M + B + TD	-8.707×10^{1}	1.540×10^{1}	-2.773×10^{4}	3.220×10^{2}	2.324×10^{1}	3.854×10^{-1}	5.546×10^{4}	6.439×10^{2}
C + P + M + B + R + TM	-8.890×10^{1}	1.507×10^{1}	-2.773×10^{4}	3.220×10^{2}	3.574×10^{1}	7.077×10^{-1}	5.546×10^{4}	6.440×10^{2}
C + P + M + B + TM	-9.082×10^{1}	1.559×10^{1}	-2.773×10^{4}	3.221×10^{2}	2.999×10^{1}	4.649×10^{-1}	5.547×10^{4}	6.441×10^{2}

Table 5.11: Model comparison for NDD Submodels: ASD

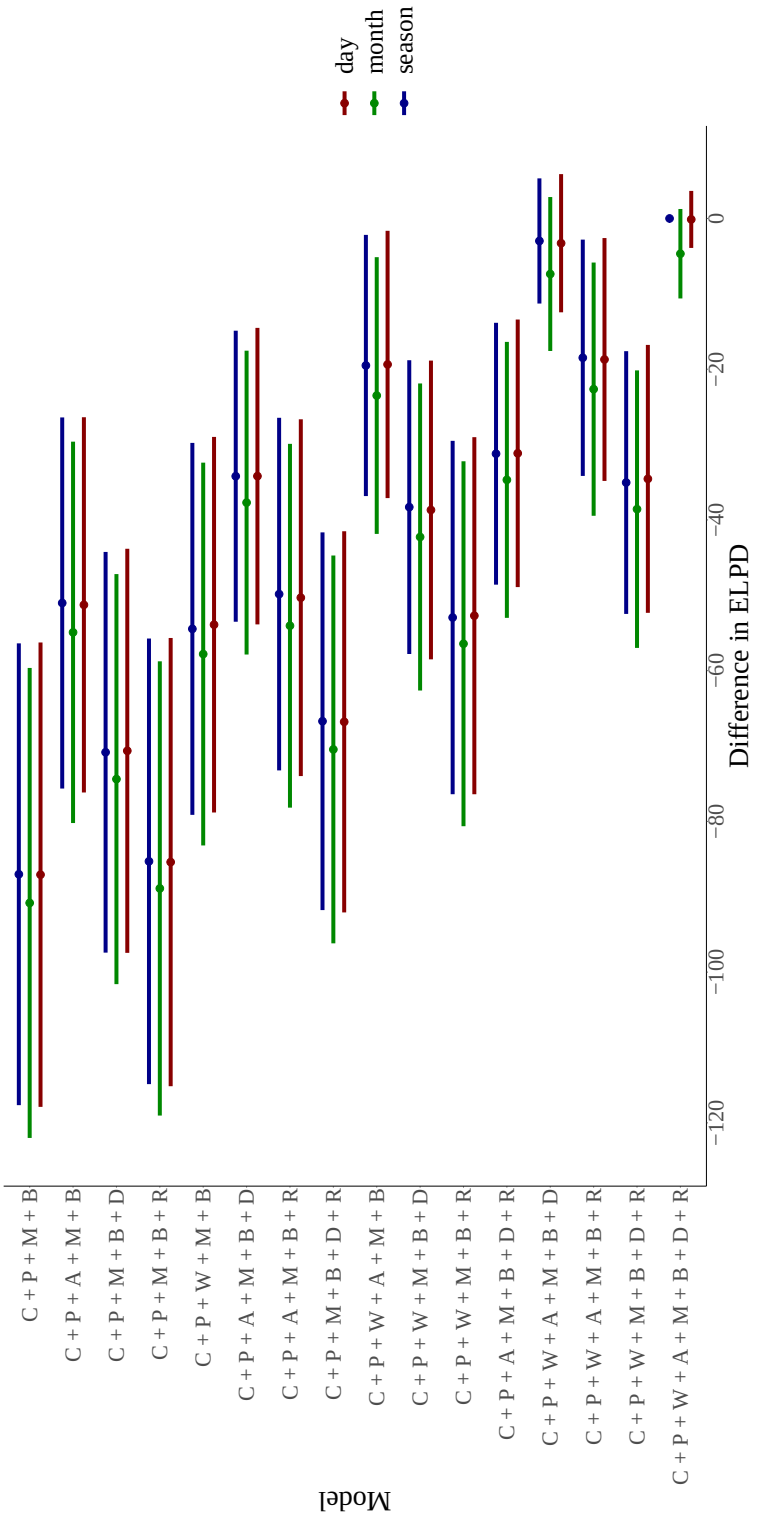


Figure 5.15: Model comparison for Neurodvelopmental submodels: ASD

model	elpd_diff	se_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
C + P + M + B + R + TD	0.000×10^{0}	0.000×10^{0}	-5.454×10^{3}	1.641×10^{2}	2.546×10^{1}	1.100×10^{0}	1.091×10^{4}	3.282×10^{2}
C + P + W + M + B + R + TD	-3.891×10^{-2}	1.806×10^{0}	-5.454×10^{3}	1.642×10^{2}	2.724×10^{1}	1.298×10^{0}	1.091×10^{4}	3.283×10^{2}
C + P + A + M + B + R + TD	-4.943×10^{-1}	1.677×10^{0}	-5.455×10^{3}	1.642×10^{2}	2.750×10^{1}	1.629×10^{0}	1.091×10^{4}	3.284×10^{2}
C + P + M + B + R + TS	-5.771×10^{-1}	1.239×10^{0}	-5.455×10^{3}	1.641×10^{2}	2.670×10^{1}	1.129×10^{0}	1.091×10^{4}	3.282×10^{2}
C + P + W + A + M + B + R + TD	-6.685×10^{-1}	2.389×10^{0}	-5.455×10^{3}	1.642×10^{2}	2.931×10^{1}	1.764×10^{0}	1.091×10^{4}	3.285×10^{2}
C + P + A + M + B + R + TS	-1.143×10^{0}	2.100×10^{0}	-5.455×10^{3}	1.642×10^{2}	2.871×10^{1}	1.613×10^{0}	1.091×10^{4}	3.284×10^{2}
C + P + W + M + B + R + TS	-1.248×10^{0}	2.179×10^{0}	-5.456×10^{3}	1.642×10^{2}	2.914×10^{1}	1.358×10^{0}	1.091×10^{4}	3.284×10^{2}
C + P + W + A + M + B + R + TS	-1.295×10^{0}	2.715×10^{0}	-5.456×10^{3}	1.642×10^{2}	3.065×10^{1}	1.798×10^{0}	1.091×10^{4}	3.285×10^{2}
C + P + M + B + D + R + TD	-2.979×10^{0}	3.587×10^{0}	-5.457×10^{3}	1.643×10^{2}	3.534×10^{1}	1.491×10^{0}	1.091×10^{4}	3.285×10^{2}
C + P + A + M + B + D + R + TD	-3.358×10^{0}	3.921×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.718×10^{1}	1.920×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + W + M + B + D + R + TD	-3.648×10^{0}	3.999×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.780×10^{1}	1.694×10^{0}	1.092×10^{4}	3.285×10^{2}
C + P + W + M + B + D + R + TS	-3.852×10^{0}	4.109×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.857×10^{1}	1.712×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + M + B + D + R + TS	-3.911×10^{0}	3.698×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.680×10^{1}	1.536×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + W + A + M + B + D + R + TD	-3.962×10^{0}	4.302×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.950×10^{1}	2.047×10^{0}	1.092×10^{4}	3.287×10^{2}
C + P + M + B + R + TM	-4.002×10^{0}	3.389×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.506×10^{1}	1.387×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + A + M + B + D + R + TS	-4.138×10^{0}	4.031×10^{0}	-5.458×10^{3}	1.643×10^{2}	3.851×10^{1}	1.923×10^{0}	1.092×10^{4}	3.287×10^{2}
C + P + A + M + B + R + TM	-4.215×10^{0}	3.739×10^{0}	-5.459×10^{3}	1.643×10^{2}	3.676×10^{1}	1.805×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + W + M + B + R + TM	-4.410×10^{0}	3.852×10^{0}	-5.459×10^{3}	1.643×10^{2}	3.719×10^{1}	1.583×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + W + A + M + B + D + R + TS	-4.555×10^{0}	4.365×10^{0}	-5.459×10^{3}	1.643×10^{2}	4.063×10^{1}	2.081×10^{0}	1.092×10^{4}	3.286×10^{2}
C + P + W + A + M + B + R + TM	-4.900×10^{0}	4.163×10^{0}	-5.459×10^{3}	1.643×10^{2}	3.915×10^{1}	1.988×10^{0}	1.092×10^{4}	3.287×10^{2}
C + P + M + B + D + TD	-6.777×10^{0}	5.894×10^{0}	-5.461×10^{3}	1.644×10^{2}	3.037×10^{1}	1.301×10^{0}	1.092×10^{4}	3.288×10^{2}
C + P + M + B + D + R + TM	-7.155×10^{0}	4.785×10^{0}	-5.461×10^{3}	1.644×10^{2}	4.508×10^{1}	1.786×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + W + M + B + D + TD	-7.290×10^{0}	6.189×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.275×10^{1}	1.518×10^{0}	1.092×10^{4}	3.288×10^{2}
C + P + A + M + B + D + TD	-7.301×10^{0}	6.106×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.231×10^{1}	1.693×10^{0}	1.092×10^{4}	3.288×10^{2}
C + P + W + M + B + D + R + TM	-7.326×10^{0}	5.111×10^{0}	-5.462×10^{3}	1.644×10^{2}	4.696×10^{1}	1.942×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + M + B + D + TS	-7.374×10^{0}	5.931×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.177×10^{1}	1.340×10^{0}	1.092×10^{4}	3.287×10^{2}
C + P + W + A + M + B + D + TD	-7.868×10^{0}	6.365×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.465×10^{1}	1.916×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + W + A + M + B + D + R + TM	-7.968×10^{0}	5.345×10^{0}	-5.462×10^{3}	1.645×10^{2}	4.901×10^{1}	2.316×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + W + M + B + D + TS	-8.038×10^{0}	6.208×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.427×10^{1}	1.559×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + A + M + B + D + TS	-8.053×10^{0}	6.133×10^{0}	-5.462×10^{3}	1.644×10^{2}	3.387×10^{1}	1.849×10^{0}	1.092×10^{4}	3.289×10^{2}
C + P + A + M + B + D + R + TM	-8.184×10^{0}	5.015×10^{0}	-5.462×10^{3}	1.645×10^{2}	4.753×10^{1}	2.178×10^{0}	1.092×10^{4}	3.290×10^{2}
C + P + W + A + M + B + D + TS	-8.618×10^{0}	6.384×10^{0}	-5.463×10^{3}	1.645×10^{2}	3.624×10^{1}	1.946×10^{0}	1.093×10^{4}	3.290×10^{2}
C + P + M + B + TD	-8.943×10^{0}	5.157×10^{0}	-5.463×10^{3}	1.644×10^{2}	2.059×10^{1}	9.131×10^{-1}	1.093×10^{4}	3.288×10^{2}
C + P + M + B + TS	-9.034×10^{0}	5.304×10^{0}	-5.463×10^{3}	1.645×10^{2}	2.180×10^{1}	9.354×10^{-1}	1.093×10^{4}	3.289×10^{2}
C + P + W + M + B + TD	-9.061×10^{0}	5.464×10^{0}	-5.463×10^{3}	1.644×10^{2}	2.264×10^{1}	1.144×10^{0}	1.093×10^{4}	3.288×10^{2}
C + P + W + M + B + TS	-9.189×10^{0}	5.586×10^{0}	-5.463×10^{3}	1.644×10^{2}	2.385×10^{1}	1.149×10^{0}	1.093×10^{4}	3.289×10^{2}
C + P + A + M + B + TD	-9.200×10^{0}	5.423×10^{0}	-5.463×10^{3}	1.645×10^{2}	2.249×10^{1}	1.457×10^{0}	1.093×10^{4}	3.289×10^{2}
C + P + A + M + B + TS	-9.494×10^{0}	5.549×10^{0}	-5.464×10^{3}	1.645×10^{2}	2.384×10^{1}	1.503×10^{0}	1.093×10^{4}	3.290×10^{2}
C + P + W + A + M + B + TS	-9.530×10^{0}	5.823×10^{0}	-5.464×10^{3}	1.645×10^{2}	2.574×10^{1}	1.643×10^{0}	1.093×10^{4}	3.290×10^{2}
C + P + W + A + M + B + TD	-9.703×10^{0}	5.705×10^{0}	-5.464×10^{3}	1.645×10^{2}	2.479×10^{1}	1.609×10^{0}	1.093×10^{4}	3.290×10^{2}
C + P + W + M + B + D + TM	-1.094×10^{1}	6.841×10^{0}	-5.465×10^{3}	1.645×10^{2}	4.200×10^{1}	1.763×10^{0}	1.093×10^{4}	3.290×10^{2}
C + P + M + B + D + TM	-1.105×10^{1}	6.625×10^{0}	-5.465×10^{3}	1.645×10^{2}	4.025×10^{1}	1.601×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + W + A + M + B + D + TM	-1.165×10^{1}	7.018×10^{0}	-5.466×10^{3}	1.646×10^{2}	4.411×10^{1}	2.158×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + A + M + B + D + TM	-1.174×10^{1}	6.791×10^{0}	-5.466×10^{3}	1.645×10^{2}	4.240×10^{1}	2.032×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + W + M + B + TM	-1.228×10^{1}	6.335×10^{0}	-5.467×10^{3}	1.645×10^{2}	3.167×10^{1}	1.378×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + M + B + TM	-1.254×10^{1}	6.070×10^{0}	-5.467×10^{3}	1.646×10^{2}	3.005×10^{1}	1.197×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + W + A + M + B + TM	-1.258×10^{1}	6.513×10^{0}	-5.467×10^{3}	1.646×10^{2}	3.352×10^{1}	1.811×10^{0}	1.093×10^{4}	3.291×10^{2}
C + P + A + M + B + TM	-1.277×10^{1}	6.288×10^{0}	-5.467×10^{3}	1.646×10^{2}	3.188×10^{1}	1.659×10^{0}	1.093×10^{4}	3.292×10^{2}

Table 5.12: Model comparison for NDD Submodels: Learning Difficulties

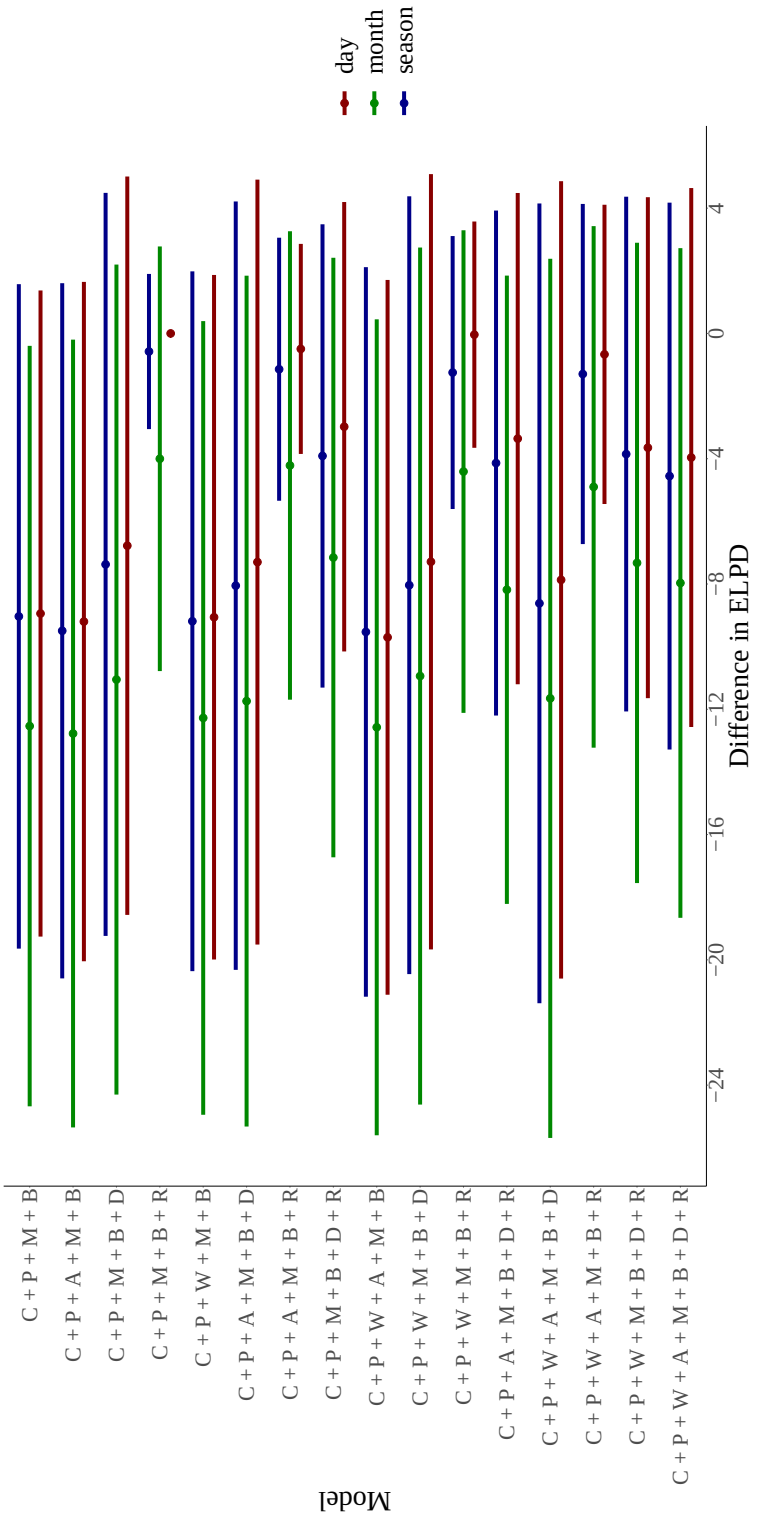


Figure 5.16: Model comparison for Neurodvelopmental submodels: Learning Difficulties

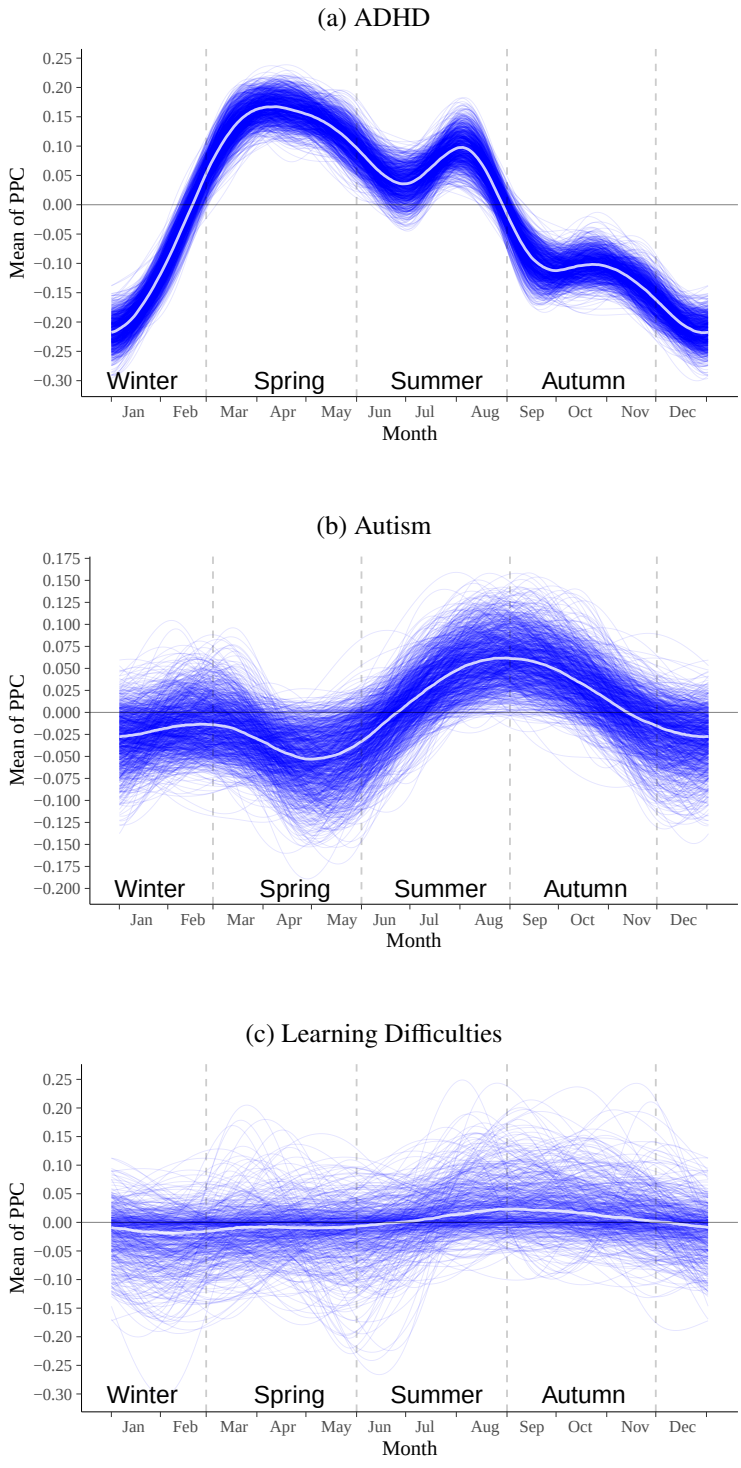


Figure 5.17: Mean of the posterior predictive distribution and randomly-sampled 1000 posterior draws of the smoothing terms for the NDD submodels

CHAPTER 6

MOLECULAR FINGERPRINTS ARE A SIMPLE YET EFFECTIVE SOLUTION TO THE DRUG–DRUG INTERACTION PROBLEM

An earlier version of this chapter was presented as a Spotlight at the 2022 ICML Workshop on Computational Biology (Long et al., 2022).

6.1 Introduction

Drug–drug interactions (DDIs), or non-additive action of multiple co-administered medications, also known as polypharmacy problem, is a significant source of adverse medical outcomes (Zitnik et al., 2018). The major difficulty in identifying and anticipating DDIs is the combinatorial explosion of potential drug interactions that renders clinical testing of all pairs impractical. Machine learning models trained on expert-curated databases and electronic health records have since provided a tractable solution to identifying potential drug interactions.

Earlier work on DDI (Gottlieb et al., 2012; Vilar et al., 2012, 2014; Cheng and Zhao, 2014) primarily computed similarity measures between a combination of various handpicked chemical properties and structural fingerprints of drugs to predict new drug-drug interactions. One drawback to these DDI approaches is that they use fixed feature representations of the drugs that may not be optimal for predicting drug interactions.

More recent machine learning methods for DDI have since focused on deep neural networks for the task. One popular approach extracts DDI information from text data and applies LSTM (Sahu and Anand, 2018) and attention based architectures (Zheng et al., 2017) on the corpus to predict drug interactions. Zhao et al. (2016) applies 1D convolutional neural networks on sentences for the task. More recent approaches rely on graph neural networks (GNNs) (Battaglia et al., 2018) automatically learning the molecular representation of the drugs to produce the interaction prediction. Zitnik et al. (2018) proposed one of the earlier approaches to use GNNs as encoders to featurize

representation of molecules. Similar approaches for learning graph representations for the drugs was proposed by Yin et al. (2022) and Wang et al. (2022). They use a concatenation of handpicked chemical features and the learned features from applying graph neural network layers (Veličković et al., 2018; Kipf and Welling, 2017) layers on the molecular structure of the pair of drugs.

Model suggested by Nyamabo et al. (2021) applies multiple graph attention layers on the individual drugs and aggregates their intermediate graph representations with co-attention to produce the output class predictions. Other works cast the problem as a knowledge graph learning problem Lin et al. (2020). The advantage of these GNN approaches is that they are able to learn feature representations for the drugs that are adapted to the classification task at hand without the need for explicit manual feature selection. While the end-to-end GNN approach has proven to be effective on a variety of DDI benchmark tasks, we believe that fingerprint based approaches still have merit. In our experiments, we show that neural networks that operate on just 2D Morgan fingerprints outperform GNN models and achieve these improved results with simpler model architectures that were much faster than the GNN baselines. Moreover, a central concern about neural networks is that they are opaque and hard to interpret. The benefit to fingerprint based neural networks is that we retain the flexibility of a neural network that can learn useful nonlinear combinations of our features while still retaining a level of intepretability by virtue of using fingerprints as our input representations.

Our contributions in this work are two-fold:

1. We curated a new dataset that augments the DrugBank dataset with negative examples derived from a large commercial insurance dataset;

2. We demonstrated that the use of Morgan fingerprints with simple neural network architectures achieved SOTA performance, outperforming GNN architectures.

6.2 Problem Setup

To formalize our problem, let D denote the set of drugs considered in our work, encoded by their DrugBank IDs. Our dataset is then a collection of triplets: $\left\{\left(d_i^{(1)}, d_i^{(2)}, l_i\right)\right\}_{i=1}^{n}$, where $l_i \in \{1, 2, \ldots, K\}$ denotes the interaction between the pair of drugs $\left(d_i^{(1)}, d_i^{(2)}\right) \in D \times D$. We treat the learning task as a multi-class classification over K classes. Hence, our objective is to learn a classification function

$$f : D \times D \to \{1, 2, \ldots, K\} \tag{6.1}$$

In the DDI literature, we also commonly see the problem cast as a binary classification task over the triplets $\left(d_i^{(1)}, d_i^{(2)}, l_i\right)$ where one predicts if the given label is in fact a true observed interaction between the pair of drugs. In both the multiclass classification setting and the binary classification setting, practitioners occasionally augment the observed DDI data with negative examples to make the models more robust. This is because we care not just about the presence of interaction but also the absence thereof when screen drugs for "safe" combinations.

6.3 Dataset

The positive DDI examples came from the DrugBank dataset (Wishart et al., 2018), while the negative examples came from IBM MarketScan (IBM Watson Health, 2019), a commercial insurance database containing > 180 million unique subjects in the United States. The prescription (RX) records from MarketScan contained information on the identity of the drugs encoded with National Drug Codes (NDCs), start date and end date. Using the dates we identified potential interactions as any overlap in the intervals of validity between any two drugs using an interval tree (Cormen et al., 2022, §17.3). Each row of data in this dataset corresponds to a pair of drugs, represented by their DrugBank IDs and the associated target value, which is a label indicating the type of the resulting interaction between the two drugs. The DrugBank IDs can be mapped to the corresponding

SMILES strings, from which we constructed the Morgan fingerprints (FPs) (Morgan, 1965) of the drugs using RDKit (Landrum et al., 2006). For more information about the dataset, see Table 6.1.

# Drugs with SMILES	3,516
# DDI Pairs	1,457,198
# Interaction Types	256

Table 6.1: Dataset details

6.4 Model Architectures

We tested a variety of models on our new drug drug interaction dataset. Our models follow the same general architecture:

1. map the SMILES string to a numerical object e.g. a binary fingerprint or matrices containing atom features and connectivity information)
2. encode the pair of input drugs with an encoder network (e.g.: a multilayer perceptron (MLP) or a graph neural network (GNN))
3. construct a pair drug representation by concatenating or adding the two individual drug representations
4. feed the pair drug representation through a final multilayer perceptron to produce a class output.

We also experimented with variants of the above architecture where the two fingerprint vectors were aggregated before being fed through an MLP instead of after step 2 above.

6.4.1 Molecular Fingerprint Model

We first generated the Morgan fingerprints, which is a kind of extended-connectivity fingerprint, from the SMILES strings with RDKit using a radius of 2 (Rogers and Hahn, 2010). The fingerprints

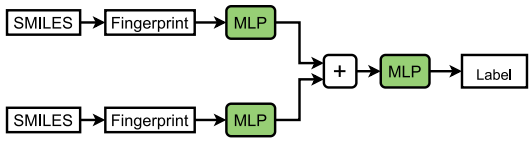


Figure 6.1: Fingerprint based neural network for predicting the interaction between two input drugs

are 2048 dimensional binary vectors which are passed through a multilayer perceptron (MLP). The resulting representation vectors are aggregated (either through concatenation or summation) to form a drug-pair representation that is sent through a final multilayer perceptron (MLP) to produce an output label. We applied the LeakyReLU (Zhang et al., 2017) activation function in the MLP layers.

6.4.2 Graph Neural Network Based Models

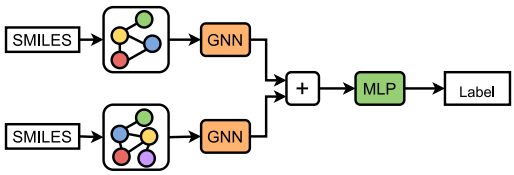


Figure 6.2: Graph neural network based architecture for predicting the interaction between two input drugs.

In the graph neural network based models, we extract atom features and atom connectivity information from the SMILES strings of the pair of drugs. The atom features, edge features and connectivity are then used as input to a graph neural network (GNN) applies multiple rounds of message passing (Gilmer et al., 2017) to produce a molecular encoding. As in the FP network, the drug feature vectors are aggregated and then passed through a final MLP to produce the predicted output label.

6.4.3 Attention-Based Models

In addition, we also compared our FP model against co-attention based model SSI-DDI-v2 that mostly follows the architecture proposed in Nyamabo et al. (2021). This model takes the two

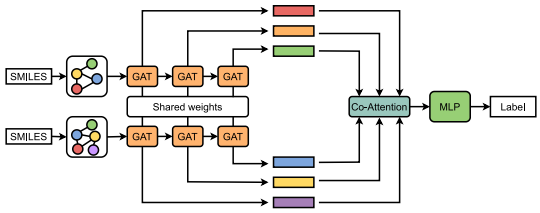


Figure 6.3: Co-attention architecture based off Nyamabo et al. (2021)'s model with intermediate graph attention layers for predicting the interaction between two input drugs. The number of GAT layers is a tunable hyperparameter.

molecular graphs with atom features from each drug in a pair as input and uses the Graph Attention Network (GAT) (Veličković et al., 2018; Brody et al., 2021) autoregressively to generate multiple layers of hidden representations for each drug separately, which are subsequently passed to an additive co-attention layer that outputs attention scores for re-weighting the hidden representations (cf. Figure 6.3) during aggregation. The final aggregated hidden representation of the drug pair is then sent to an MLP to produce the predicted output label. Other models making use of self- and/or co-attention include Lin et al. (2022); Nyamabo et al. (2022); Pang et al. (2022).

6.5 Experiments

We evaluated each of the models on our dataset using $80/10/10$ training/validation/test split. Each of the networks produced a softmax probability distribution over all the DDI classes for predicting the interaction between drug pairs. We used the RAdam optimizer (Liu et al., 2019) to minimize the cross-entropy loss over each mini-batch, with an early stopping $\Delta = 5 \times 10^{-3}$ and a tolerance of 4 epochs on validation loss. All models were implemented using PyTorch (Paszke et al., 2019) and Pytorch Geometric (Fey and Lenssen, 2019) and experiment logging was done with PyTorch Lightning.

6.5.1 Hyperparameter Tuning

We manually tuned the hyperparameters of all models trained. Table 6.2 details the various hyperparameters we used and experimented with.

Parameter	Values
AdamW Learning Rate	$\{10^{-5}, 10^{-4}, 10^{-3}\}$
AdamW Weight Decay	$\{0\}$
MLP Layers	$\{3, 4, 5\}$
GNN Layers	$\{3, 4, 5\}$
Minibatch size	{128, 256, 512}

Table 6.2: Hyperparameter Values

6.6 Results and Discussion

Table 6.4 shows the results of our experiments. The FP model achieved higher performance on all of the reported metrics and converged in fewer training epochs with less time per epoch than the graph neural network baseline methods (Table 6.3). Thus, we now focus on more detailed comparison between the FP model and SSI-DDI-v2. Figure 6.4 shows the values of the metrics among classes that had at least 100 instances in the test dataset (56 classes), with decreasing number of instances per class. Overall, FP and SSI-DDI-v2 had similar performance profile *across* the classes, achieving full accuracy on class 226, corresponding to the interaction string

```
The risk or severity of tendinopathy can be increased when Drug1 is
combined with Drug2.
```

Figure 6.5 shows the difference in accuracy (6.5a) and weighted-F_1 score (6.5b) between FP (reference) and SSI-DDI-v2. The largest *positive* difference was found in class 193, corresponding to the interaction string

```
The risk or severity of myopathy, rhabdomyolysis, and myoglobinuria
can be increased when Drug1 is combined with Drug2.
```

On the other hand, the largest *negative* difference was found in class 95, corresponding to the interaction string

```
The protein binding of Drug1 can be decreased when combined with Drug2.
```

6.7 Additional Tables and Figures

Model	Training epochs	Avg. time/epoch (min)
FP	**15**	**1.37**
GCNConv (Kipf and Welling, 2017)	34	4.09
GATConv (Veličković et al., 2018)	27	5.00
GATv2Conv (Brody et al., 2021)	19	5.42
SSI-DDI (Nyamabo et al., 2021)	30	7.83
SSI-DDI-v2 (Nyamabo et al., 2021)	25	8.93

Table 6.3: Training information: number of epochs until convergence and time per epoch. SSI-DDI uses GATConv by default. SSI-DDI-v2 uses GATv2Conv instead. All GNN models used 4 layers.

Model	Accuracy	Macro-F_1	Weighted-F_1	AUROC
FP	**0.9615**	**0.9213**	**0.9612**	**0.9989**
GCNConv	0.8694	0.7144	0.8603	0.9793
GATConv	0.7958	0.6019	0.7763	0.9650
GATv2Conv	0.8696	0.7107	0.8622	0.9880
SSI-DDI	0.9422	0.8903	0.9415	0.9915
SSI-DDI-v2	0.9491	0.9047	0.9488	0.9923

Table 6.4: Test results on a hold-out set.

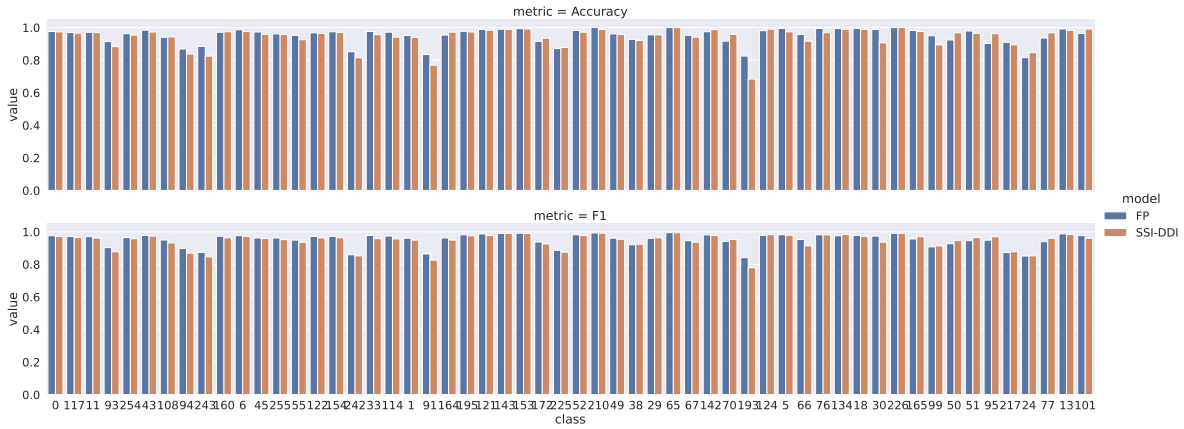
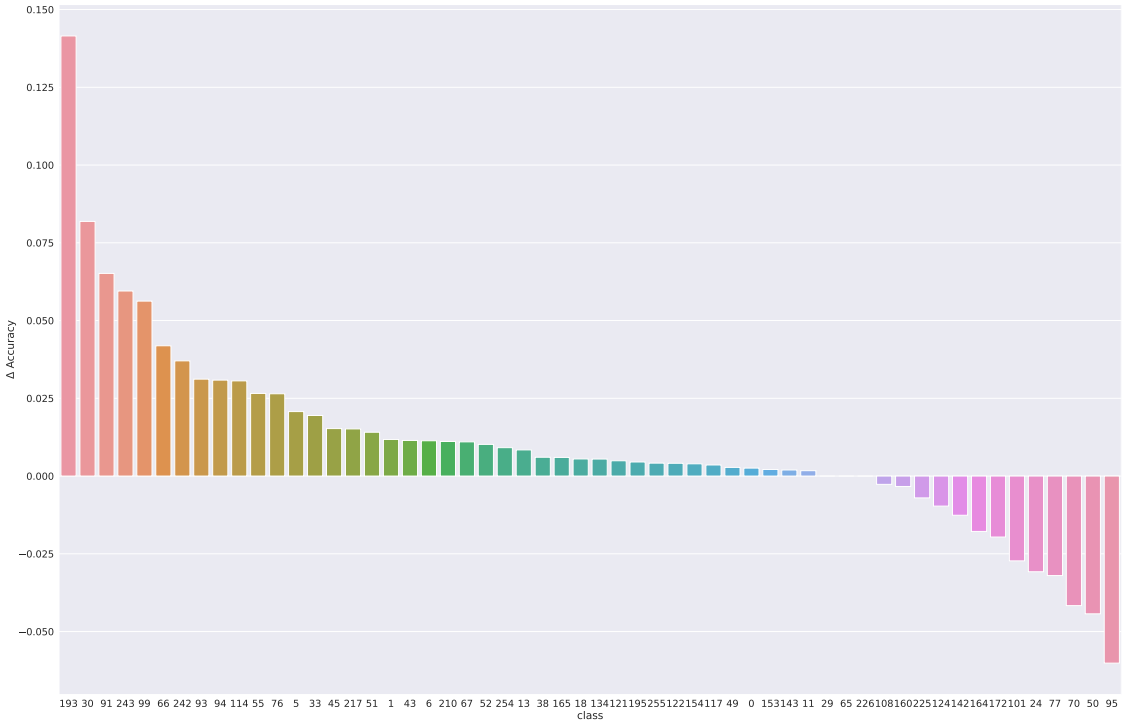
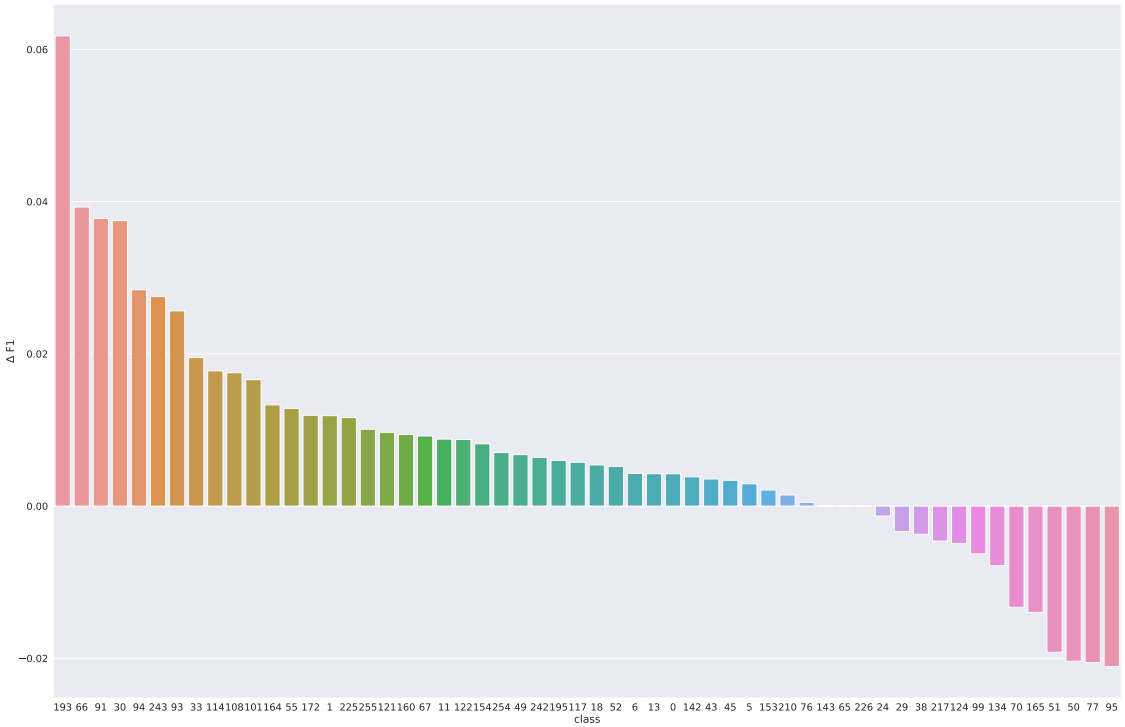


Figure 6.4: Direct comparison of metrics by class between FP and SSI-DDI-v2 for classes with at least 100 instances in the test dataset, sorted by decreasing number of instances in each class



(a) Accuracy



(b) Weighted-F_1

Figure 6.5: Differences (FP vs. SSI-DDI-v2) in metrics by class for classes with at least 100 instances in the test dataset

CHAPTER 7

CONCLUSION

In this study, we harnessed the statistical power of IBM MarketScan, a very large commercial insurance claims dataset, to study the effect of a diverse range of health outcomes during the early life of a newborn. For the SRB study, we probed a diverse range of associations with exogenous factors and used these results to further challenge popular theories in sexual selection. For the immune system disorders and NDDs, we have situated our results within the burgeoning field of microbiota-gut-brain (MGB) axis, noting that early life and prenatal maternal exposure to infections and anti-infective prescriptions likely compromised the newborn's immune system, leading to higher risks of diseases of both the immune system and the nervous system.

In the sex ratio at birth study, we identified a large number of associations between environmental pollution and altered SRB, which, if verified in other cohorts and geographical locations, could be used for public health surveillance. Moreover, as the results did not form any clear pattern with respect to the adversity of the exogenous factors, suggesting that adaptive theories that have enjoyed wide popularity, such as the Trivers–Willard hypothesis (TWH) (Trivers and Willard, 1973), were unlikely to hold for human populations. Indeed, the results would even threaten the status of an even better established theory in sexual selection — the Düsing–Fisher theory of equal investment: that in an population with unbalanced sex ratio, the rarer sex will have on average higher reproductive value than the more frequent sex (West, 2009, Ch. 2). The predicament of these two theories were amplified by the main findings of (Zietsch et al., 2020) that used a Swedish cohort from 1932–2014, showing that the sex ratio of offspring was not heritable. Therefore, the Düsing–Fisher theory could not have operated on SRB since it is unable to explain how sex ratio became non-heritable even if it had once been heritable at some point in the evolutionary history of humans. While unlike in (Zietsch et al., 2020), we did not have access to any actual population registers, the cohort used in that study would be highly similar at least to modern Western birth cohorts, especially with respect to mechanisms for sexual selection. Finally, a recent simulation study has shown that whereas the

population sex ratio reached parity at equilibrium, the sex ratio was not heritable (Harper et al., 2023). Therefore, together with those from other recent studies, our results support the conclusion that the Düsing–Fisher theory was also not appropriate for explaining sexual selection in humans and that SRB may be determined by random Mendelian segregation as well as possible environmental factors. A related issue is how results from SRB studies in other animal species may be relevant for human SRB. If TWH and other adaptive theories turn out to not operate on contemporary human population, one should not expect the situation of other animals to generalize to any population (Kokko and Jennions, 2008).

On the other hand, the neurodevelopemental study showed that the being of male sex, C-section mode of delivery, abnormal birth weights, as well as an array of immune-system related diagnosis and anti-ininfective prescriptions both during pregnancy and in the early life, contributed to elevated risks of NDDs. In addition, for ADHD specifically, positive associations have also been identified for elevated levels of $PM_{2.5}$ — all of these we suggest may be due to compromised immune system during pregnancy and early in life. However, based on previous studies, we caution against a direct causal interpretation for antiinfectives and $PM_{2.5}$ as prior twins and siblings studies did not detect any statistically significant associations. One intriguing result was that being born in the spring and summer also contributed to higher OR of being diagnosed of ADHD. While the peak located near September 1 can be explained by the differential level of socialization of the youngest in the cohort, we found the peak in spring surprising.

Our overall approach also had a number of limitations. While the large sample size of the MarketScan dataset was surely an advantage, it is not without potential biases. First, the dataset only included commercially-insured enrollees, most of whom obtained the policies via employment. These would exclude, among others, unemployed and/or uninsured individuals who generally were of lower socioeconomic status or the elderly on medicare (age $\geqslant$ 65), but the latter group was not relevant for our studies. These subpopulations likely had higher incidences of the kind of health events that we have studied. Moreover, the data were pooled over a large number of insurance

companies, with the billing practices of larger companies dominating those of smaller ones, all of which may have divergent practices, so the same code may or may not represent the same underlying diagnosis or phenotype. More importantly, unlike electronic health records (EHRs) maintained by hospitals, we did not have access to free-text diagnosis information, which may contain information on other diagnoses or phenotypes that were present but not entered into the system for insurance billing purposes and were thus left out of the MarketScan dataset altogether. Future studies should focus on datasets that provide free texts comments by the healthcare providers in addition to the diagnoses codes. Finally, a more fine-grained modelling should include the temporal dynamics of the immune system- and neurodevelpmental phenotypes, which could take the form of a survival analysis that models the time until the NDD diagnosis or a joint longitudinal analysis.

REFERENCES

Aghaei, M., Janjani, H., Yousefian, F., Jamal, A., and Yunesian, M. (2019). Association between ambient gaseous and particulate air pollutants and attention deficit hyperactivity disorder (adhd) in children; a systematic review. *Environmental research*, 173:135–156.

Agresti, A. (2015). *Foundations of linear and generalized linear models*. John Wiley & Sons.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Amico, M. and Van Keilegom, I. (2018). Cure models in survival analysis. *Annual Review of Statistics and Its Application*, 5:311–342.

Arnold, S. J. and Duvall, D. (1994). Animal mating systems: a synthesis based on selection theory. *The American Naturalist*, 143(2):317–348.

Axelsson, P. B., Clausen, T. D., Petersen, A. H., Hageman, I., Pinborg, A., Kessing, L. V., Bergholt, T., Rasmussen, S. C., Keiding, N., and Løkkegaard, E. C. L. (2019a). Investigating the effects of cesarean delivery and antibiotic use in early childhood on risk of later attention deficit hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, 60(2):151–159.

Axelsson, P. B., Clausen, T. D., Petersen, A. H., Hageman, I., Pinborg, A., Kessing, L. V., Bergholt, T., Rasmussen, S. C., Keiding, N., and Løkkegaard, E. C. L. (2019b). Relation between infant microbiota and autism?: results from a national cohort sibling design study. *Epidemiology*, 30(1):52–60.

Bach, J.-F. (2002). The effect of infections on susceptibility to autoimmune and allergic diseases. *New England journal of medicine*, 347(12):911–920.

Bateman, A. J. (1948). Intra-sexual selection in drosophila. *Heredity*, 2(3):349–368.

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., et al. (2018). Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*.

Bauman, M. D., Iosif, A.-M., Smith, S. E., Bregere, C., Amaral, D. G., and Patterson, P. H. (2014). Activation of the maternal immune system during pregnancy alters behavioral development of rhesus monkey offspring. *Biological psychiatry*, 75(4):332–341.

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

Beversdorf, D. Q., Stevens, H. E., and Jones, K. L. (2018). Prenatal stress, maternal immune dysregulation, and their association with autism spectrum disorders. *Current psychiatry reports*, 20(9):1–12.

Bird, A. (2017). Inference to the best explanation, bayesianism, and knowledge. *Best explanations: new essays on inference to the best explanation*, pages 97–120.

Boklage, C. E. (2005). The epigenetic environment: secondary sex ratio depends on differential survival in embryogenesis. *Human Reproduction*, 20(3):583–587.

Bowker, G. C. and Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT press.

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, 71(356):791–799.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31):334–355.

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., Scott, S. L., et al. (2015). Inferring causal impact using bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1):247–274.

Brodin, P., Jojic, V., Gao, T., Bhattacharya, S., Angel, C. J. L., Furman, D., Shen-Orr, S., Dekker, C. L., Swan, G. E., Butte, A. J., et al. (2015). Variation in the human immune system is largely driven by non-heritable influences. *Cell*, 160(1-2):37–47.

Brody, S., Alon, U., and Yahav, E. (2021). How attentive are graph attention networks? *arXiv preprint arXiv:2105.14491*.

Brown, G. R. and Silk, J. B. (2002). Reconsidering the null hypothesis: is maternal rank associated with birth sex ratios in primate groups? *Proceedings of the National Academy of Sciences*, 99(17):11252–11255.

Brown, R. L. and Clarke, T. B. (2017). The regulation of host defences to infection by the microbiota. *Immunology*, 150(1):1–6.

Bruckner, T. A. and Catalano, R. (2018). Selection in utero and population health: theory and typology of research. *SSM-population health*, 5:101–113.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28.

Bürkner, P.-C., Gabry, J., Kay, M., and Vehtari, A. (2021). posterior: Tools for working with posterior distributions. R package version 1.1.0.

Canali, S. and Leonelli, S. (2022). Reframing the environment in data-intensive health sciences. *Studies in History and Philosophy of Science*, 93:203–214.

Cantwell, D. P. and Baker, L. (1991). Association between attention deficit-hyperactivity disorder and learning disorders. *Journal of learning disabilities*, 24(2):88–95.

Carlsson, T., Molander, F., Taylor, M. J., Jonsson, U., and Bölte, S. (2021). Early environmental risk factors for neurodevelopmental disorders–a systematic review of twin and sibling studies. *Development and Psychopathology*, 33(4):1448–1495.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2017). Stan: a probabilistic programming language. *Grantee Submission*, 76(1):1–32.

Casanova, J.-L. and Abel, L. (2007). Human genetics of infectious diseases: a unified theory. *The EMBO journal*, 26(4):915–922.

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

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

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

Catalano, R., Gemmill, A., Casey, J., Karasek, D., Stewart, H., and Saxton, K. (2018). Separating the bruce and trivers-willard effects in theory and in human data. *American Journal of Human Biology*, 30(2):e23074.

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

Cattalini, M., Soliani, M., Caparello, M. C., and Cimaz, R. (2019). Sex differences in pediatric rheumatology. *Clinical reviews in allergy & immunology*, 56(3):293–307.

Champagne-Jorgensen, K., Kunze, W. A., Forsythe, P., Bienenstock, J., and Neufeld, K.-A. M. (2019). Antibiotics and the nervous system: More than just the microbes? *Brain, behavior, and immunity*, 77:7–15.

Cheng, F. and Zhao, Z. (2014). Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *Journal of the American Medical Informatics Association*, 21(e2):e278–e286.

Choi, G. B., Yim, Y. S., Wong, H., Kim, S., Kim, H., Kim, S. V., Hoeffer, C. A., Littman, D. R., and Huh, J. R. (2016). The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. *Science*, 351(6276):933–939.

Chu, D. M., Ma, J., Prince, A. L., Antony, K. M., Seferovic, M. D., and Aagaard, K. M. (2017). Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. *Nature medicine*, 23(3):314–326.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). STL: A seasonal-trend decomposition. *Journal of official statistics*, 6(1):3–73.

Collado, M. C., Cernada, M., Neu, J., Pérez-Martínez, G., Gormaz, M., and Vento, M. (2015). Factors influencing gastrointestinal tract and microbiota immune interaction in preterm infants. *Pediatric research*, 77(6):726–731.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. (2022). *Introduction to Algorithms, 4th edition*. MIT press.

Cryan, J. F., O'Riordan, K. J., Cowan, C. S., Sandhu, K. V., Bastiaanssen, T. F., Boehme, M., Codagnone, M. G., Cussotto, S., Fulling, C., Golubeva, A. V., et al. (2019). The microbiota-gut-brain axis. *Physiological reviews*.

Curran, E. A., Khashan, A. S., Dalman, C., Kenny, L. C., Cryan, J. F., Dinan, T. G., and Kearney, P. M. (2016). Obstetric mode of delivery and attention-deficit/hyperactivity disorder: a sibling-matched study. *International journal of epidemiology*, 45(2):532–542.

De Lisle, S. P. (2019). Understanding the evolution of ecological sex differences: Integrating character displacement and the darwin-bateman paradigm. *Evolution Letters*, 3(5):434–447.

Dehner, C., Fine, R., and Kriegel, M. A. (2019). The microbiome in systemic autoimmune disease–mechanistic insights from recent studies. *Current opinion in rheumatology*, 31(2):201.

DiPietro, J. A. and Voegtline, K. M. (2017). The gestational foundation of sex differences in development and vulnerability. *Neuroscience*, 342:4–20.

Dixson, B. J., Haywood, J., Lester, P. J., and Ormsby, D. K. (2011). Whatever the weather: ambient temperature does not influence the proportion of males born in new zealand. *PloS one*, 6(9):e25064.

Douhard, M. (2017). Offspring sex ratio in mammals and the trivers-willard hypothesis: In pursuit of unambiguous evidence. *BioEssays*, 39(9):1700043.

Douven, I. (2021). Abduction. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition.

Douven, I. (2022). *The art of abduction*. MIT Press.

Dupré, J. (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Harvard University Press.

Dupré, J. (2012). *Processes of life: Essays in the philosophy of biology*. Oxford University Press.

Dupré, J. (2020). Processes within processes: A dynamic account of living beings and its implications for understanding the human individual. In Meincke, A. S. and Dupré, J., editors, *Biological Identity: Perspectives from Metaphysics and the Philosophy of Biology*, pages 149–166. Routledge.

Dupré, J. and Leonelli, S. (2022). Process epistemology in the covid-19 era: rethinking the research process to avoid dangerous forms of reification. *European Journal for Philosophy of Science*, 12(1):20.

Emilsson, L., Lindahl, B., Köster, M., Lambe, M., and Ludvigsson, J. F. (2015). Review of 103 swedish healthcare quality registries. *Journal of internal medicine*, 277(1):94–136.

Esaiassen, E., Fjalstad, J. W., Juvet, L. K., van den Anker, J. N., and Klingenberg, C. (2017). Antibiotic exposure in neonates and early adverse outcomes: a systematic review and meta-analysis. *Journal of antimicrobial chemotherapy*, 72(7):1858–1870.

Ferguson, T. S. (2017). *A course in large sample theory*. Routledge.

Fey, M. and Lenssen, J. E. (2019). Fast graph representation learning with PyTorch Geometric. In *ICLR Workshop on Representation Learning on Graphs and Manifolds*.

Fine, C. (2017). *Testosterone Rex: Myths of sex, science, and society*. WW Norton & Company.

Finn, S. (2023). The mereotopology of pregnancy. In *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, page jhad017. Oxford University Press US.

Fisher, R. A. (1930). *The genetical theory of natural selection*. The Claredon Press.

Floyd, D. N., Langham, S., Séverac, H. C., and Levesque, B. G. (2015). The economic and quality-of-life burden of crohn's disease in europe and the united states, 2000 to 2013: a systematic review. *Digestive diseases and sciences*, 60(2):299–312.

Fritzsche, K. and Arnqvist, G. (2013). Homage to bateman: sex roles predict sex differences in sexual selection. *Evolution*, 67(7):1926–1936.

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

Fukuda, M., Fukuda, K., Shimizu, T., Nobunaga, M., Mamsen, L. S., and Andersen, C. Y. (2014). Climate change is associated with male: female ratios of fetal deaths and newborn infants in japan. *Fertility and sterility*, 102(5):1364–1370.

Fung, T. C. (2020). The microbiota-immune axis as a central mediator of gut-brain communication. *Neurobiology of disease*, 136:104714.

Gabry, J. and Češnovar, R. (2020). cmdstanr: R interface to'cmdstan'.

Gabry, J. and Mahr, T. (2021). bayesplot: Plotting for bayesian models. R package version 1.8.1.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization in bayesian workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(2):389–402.

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

Gelman, A. and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*, 348:1–17.

Gelman, A. and Shalizi, C. R. (2013). Philosophy and the practice of bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38.

Gelman, A., Vehtari, A., Simpson, D., Margossian, C. C., Carpenter, B., Yao, Y., Kennedy, L., Gabry, J., Bürkner, P.-C., and Modrák, M. (2020). Bayesian workflow. *arXiv preprint arXiv:2011.01808*.

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

Gialluisi, A., Andlauer, T. F., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., Ludwig, K. U., Czamara, D., Pourcain, B. S., Honbolygó, F., et al. (2021). Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Molecular psychiatry*, 26(7):3004–3017.

Gigerenzer, G. (2018). Statistical rituals: The replication delusion and how we got there. *Advances in Methods and Practices in Psychological Science*, 1(2):198–218.

Gilbert, S. F. (2014). A holobiont birth narrative: the epigenetic transmission of the human microbiome. *Frontiers in genetics*, 5:282.

Gilbert, S. F., Bosch, T. C., and Ledón-Rettig, C. (2015). Eco-evo-devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nature Reviews Genetics*, 16(10):611–622.

Gilbert, S. F., Sapp, J., and Tauber, A. I. (2012). A symbiotic view of life: we have never been individuals. *The Quarterly review of biology*, 87(4):325–341.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. (2017). Neural message passing for quantum chemistry. In *International conference on machine learning*, pages 1263–1272. PMLR.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society*, pages 1293–1301.

Gomez-Lopez, N., StLouis, D., Lehr, M. A., Sanchez-Rodriguez, E. N., and Arenas-Hernandez, M. (2014). Immune cells in term and preterm labor. *Cellular & molecular immunology*, 11(6):571–581.

Gottlieb, A., Stein, G. Y., Oron, Y., Ruppin, E., and Sharan, R. (2012). Indi: a computational framework for inferring drug interactions and their associated recommendations. *Molecular systems biology*, 8(1):592.

Gowaty, P. A. (2015). Standing on darwin's shoulders: the nature of selection hypotheses. In *Current perspectives on sexual selection*, pages 103–118. Springer.

Gowaty, P. A. and Hubbell, S. P. (2009). Reproductive decisions under ecological constraints: it's about time. *Proceedings of the National Academy of Sciences*, 106(Supplement 1):10017–10024.

Gowaty, P. A., Kim, Y.-K., and Anderson, W. W. (2012). No evidence of sexual selection in a repetition of bateman's classic study of drosophila melanogaster. *Proceedings of the National Academy of Sciences*, 109(29):11740–11745.

Gómez-Vallejo, S., Leoni, M., Ronald, A., Colvert, E., Happé, F., and Bolton, P. (2021). Autism spectrum disorder and obstetric optimality: a twin study and meta-analysis of sibling studies. *Journal of Child Psychology and Psychiatry*, 62(11):1353–1362.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.

Grant, V. J. (2003). The maternal dominance hypothesis: questioning trivers and willard. *Evolutionary Psychology*, 1(1):147470490300100106.

Grant, V. J. (2007). Could maternal testosterone levels govern mammalian sex ratio deviations? *Journal of Theoretical Biology*, 246(4):708–719.

Grant, V. J. and Chamley, L. W. (2010). Can mammalian mothers influence the sex of their offspring peri-conceptually? *Reproduction*, 140(3):425–433.

Griffiths, J., Jenkins, P., Vargova, M., Bowler, U., Juszczak, E., King, A., Linsell, L., Murray, D., Partlett, C., Patel, M., et al. (2019). Enteral lactoferrin supplementation for very preterm infants: a randomised placebo-controlled trial. *The Lancet*, 393(10170):423–433.

Guay, A. and Pradeu, T. (2016a). *Individuals across the sciences*. Oxford University Press, USA.

Guay, A. and Pradeu, T. (2016b). To be continued: The genidentity of physical and biological processes. In Guay, A. and Pradeu, T., editors, *Individuals across the sciences*, chapter 16, pages 317–47. Oxford University Press, USA.

Han, V. X., Patel, S., Jones, H. F., Nielsen, T. C., Mohammad, S. S., Hofer, M. J., Gold, W., Brilot, F., Lain, S. J., Nassar, N., et al. (2021). Maternal acute and chronic inflammation in pregnancy is associated with common neurodevelopmental disorders: a systematic review. *Translational psychiatry*, 11(1):1–12.

Hansen, L. (2017). The Truven health MarketScan databases for life sciences researchers. *Truven Health Ananlytics IBM Watson Health*.

Harper, K., Sidari, M. J., and Zietsch, B. P. (2023). Agent-based modelling shows that operation of fisher's principle does not deplete heritability in offspring sex ratio. *bioRxiv*, pages 2023–04.

Harvey, A., Koopman, S. J., and Riani, M. (1997). The modeling and seasonal adjustment of weekly observations. *Journal of Business & Economic Statistics*, 15(3):354–368.

Hayashi, F. (2000). *Econometrics*. Princeton University Press. Section.

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

Henderickx, J. G., Zwittink, R. D., Van Lingen, R. A., Knol, J., and Belzer, C. (2019). The preterm gut microbiota: An inconspicuous challenge in nutritional neonatal care. *Frontiers in cellular and infection microbiology*, 9:85.

Hernán, M. A., Hsu, J., and Healy, B. (2019). A second chance to get causal inference right: a classification of data science tasks. *Chance*, 32(1):42–49.

Hernán, M. A. and Robins, J. M. (2020). *Causal inference*. Chapman & Hall/CRC.

Hewagama, A. and Richardson, B. (2009). The genetics and epigenetics of autoimmune diseases. *Journal of autoimmunity*, 33(1):3–11.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.

Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.

Hoquet, T. (2020a). Bateman (1948): rise and fall of a paradigm? *Animal Behaviour*, pages 223–231.

Hoquet, T. (2020b). Bateman's principles: why biology needs history and philosophy. *Animal Behaviour*, 168:e5–e9.

Hoquet, T., Bridges, W. C., and Gowaty, P. A. (2020). Bateman's data: Inconsistent with "bateman's principles". *Ecology and Evolution*, 10(19):10325–10342.

Howson, C. and Urbach, P. (1993). *Scientific reasoning : the Bayesian approach*. Open Court, Chicago, 2nd edition.

Hrdy, S. B. (1986). Empathy, polyandry, and the myth of the "coy" female. In Bleier, R., editor, *Feminist Approaches to Science*, pages 119–146. Pergamon Press.

Hsiao, E. Y., McBride, S. W., Chow, J., Mazmanian, S. K., and Patterson, P. H. (2012). Modeling an autism risk factor in mice leads to permanent immune dysregulation. *Proceedings of the National Academy of Sciences*, 109(31):12776–12781.

Hsiao, E. Y., McBride, S. W., Hsien, S., Sharon, G., Hyde, E. R., McCue, T., Codelli, J. A., Chow, J., Reisman, S. E., Petrosino, J. F., et al. (2013). Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell*, 155(7):1451–1463.

Hsu, C.-W., Tseng, P.-T., Tu, Y.-K., Lin, P.-Y., Hung, C.-F., Liang, C.-S., Hsieh, Y.-Y., Yang, Y.-H., Wang, L.-J., and Kao, H.-Y. (2021). Month of birth and mental disorders: A population-based study and validation using global meta-analysis. *Acta Psychiatrica Scandinavica*, 144(2):153–167.

Hu, W., Liu, B., Gomes, J., Zitnik, M., Liang, P., Pande, V., and Leskovec, J. (2020). Strategies for pre-training graph neural networks. In *International Conference on Learning Representations*.

Huang, K., Xiao, C., Hoang, T., Glass, L., and Sun, J. (2020). Caster: Predicting drug interactions with chemical substructure representation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 702–709.

Hubbard, R. (1990). *The politics of women's biology*. Rutgers University Press.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., and Yasmeen, F. (2019). *forecast: Forecasting functions for time series and linear models*. R package version 8.5.

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

IBM Watson Health (2019). IBM MarketScan research databases for life sciences researchers.

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer Science & Business Media.

Inkster, A. M., Fernández-Boyano, I., and Robinson, W. P. (2021). Sex differences are here to stay: Relevance to prenatal care. *Journal of Clinical Medicine*, 10(13):3000.

Jacobson, A., Yang, D., Vella, M., and Chiu, I. M. (2021). The intestinal neuro-immune axis: crosstalk between neurons, immune cells, and microbes. *Mucosal immunology*, 14(3):555–565.

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

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

James, W. H. (2013). Evolution and the variation of mammalian sex ratios at birth: reflections on trivers and willard (1973). *Journal of theoretical biology*, 334:141–148.

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

James, W. H. and Grech, V. (2018). 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*, 118:15–21.

Janicke, T., Häderer, I. K., Lajeunesse, M. J., and Anthes, N. (2016). Darwinian sex roles confirmed across the animal kingdom. *Science advances*, 2(2):e1500983.

Kamdar, S., Hutchinson, R., Laing, A., Stacey, F., Ansbro, K., Millar, M. R., Costeloe, K., Wade, W. G., Fleming, P., and Gibbons, D. L. (2020). Perinatal inflammation influences but does not arrest rapid immune development in preterm babies. *Nature communications*, 11(1):1–14.

Karlstad, Ø., Furu, K., Stoltenberg, C., Håberg, S. E., and Bakken, I. J. (2017). Adhd treatment and diagnosis in relation to children's birth month: Nationwide cohort study from norway.

Kay, M. *ggdist: Visualizations of Distributions and Uncertainty*. R package version 3.2.1.

Kay, M. *tidybayes: Tidy Data and Geoms for Bayesian Models*. R package version 3.0.3.

Keller, M. C., Nesse, R. M., and Hofferth, S. (2001). The trivers–willard hypothesis of parental investment: no effect in the contemporary united states. *Evolution and Human Behavior*, 22(5):343–360.

Khan, A., Plana-Ripoll, O., Antonsen, S., Brandt, J., Geels, C., Landecker, H., Sullivan, P. F., Pedersen, C. B., and Rzhetsky, A. (2019). Environmental pollution is associated with increased risk of psychiatric disorders in the us and denmark. *PLoS biology*, 17(8):e3000353.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kingma, E. (2019). Were you a part of your mother? *Mind*, 128(511):609–646.

Kipf, T. N. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*.

Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big data & society*, 1(1):2053951714528481.

Kitchin, R. (2022). *The Data Revolution: A Critical Analysis of Big Data, Open Data and Data Infrastructures, 2nd edition*. SAGE.

Klein, S. L. and Flanagan, K. L. (2016). Sex differences in immune responses. *Nature Reviews Immunology*, 16(10):626–638.

Kokko, H., Booksmythe, I., and Jennions, M. D. (2013). Causality and sex roles: prejudice against patterns? a reply to ah-king. *Trends in ecology & evolution*, 28(1):2–4.

Kokko, H. and Jennions, M. D. (2008). Parental investment, sexual selection and sex ratios. *Journal of evolutionary biology*, 21(4):919–948.

Kristensen, K. and Henriksen, L. (2016). Cesarean section and disease associated with immune function. *Journal of Allergy and Clinical Immunology*, 137(2):587–590.

Kruschke, J. (2015). Doing bayesian data analysis, second edition: A tutorial with r, jags, and stan.

Kruschke, J. K. (2013). Posterior predictive checks can and should be bayesian: Comment on gelman and shalizi,'philosophy and the practice of bayesian statistics'. *British Journal of Mathematical and Statistical Psychology*, 66(1):45–56.

Kruschke, J. K. (2021). Bayesian analysis reporting guidelines. *Nature Human Behaviour*, 5(10):1282–1291.

Kulaylat, A. S., Schaefer, E. W., Messaris, E., and Hollenbeak, C. S. (2019). Truven health analytics marketscan databases for clinical research in colon and rectal surgery. *Clinics in colon and rectal surgery*, 32(01):054–060.

Kuperman, A. A. and Koren, O. (2016). Antibiotic use during pregnancy: how bad is it? *BMC medicine*, 14(1):1–7.

Lai, G. C., Tan, T. G., and Pavelka, N. (2019). The mammalian mycobiome: a complex system in a dynamic relationship with the host. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 11(1):e1438.

Landrum, G. et al. (2006). Rdkit: Open-source cheminformatics.

Langdon, A., Crook, N., and Dantas, G. (2016). The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome medicine*, 8(1):1–16.

Layton, T. J., Barnett, M. L., Hicks, T. R., and Jena, A. B. (2018). Attention deficit–hyperactivity disorder and month of school enrollment. *New England Journal of Medicine*, 379(22):2122–2130.

Legendre, P. and Legendre, L. F. (2012). *Numerical ecology*, volume 24. Elsevier.

Legrand, C. (2021). *Advanced survival models*. Chapman and Hall/CRC.

Leimar, O. (1996). Life-history analysis of the trivers and willard sex-ratio problem. *Behavioral Ecology*, 7(3):316–325.

Leiser, C. L., Hanson, H. A., Sawyer, K., Steenblik, J., Al-Dulaimi, R., Madsen, T., Gibbins, K., Hotaling, J. M., Ibrahim, Y. O., VanDerslice, J. A., et al. (2019). Acute effects of air pollutants on spontaneous pregnancy loss: a case-crossover study. *Fertility and sterility*, 111(2):341–347.

Leonelli, S. (2016). *Data-Centric Biology: A Philosophical Study*. University of Chicago Press.

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

Leung, J. M., Graham, A. L., and Knowles, S. C. (2018). Parasite-microbiota interactions with the vertebrate gut: synthesis through an ecological lens. *Frontiers in microbiology*, 9:843.

Li, H., Calder, C. A., and Cressie, N. (2007). Beyond moran's i: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39(4):357–375.

Li, Y., Wang, Y., and Zhang, T. (2022). Fecal microbiota transplantation in autism spectrum disorder. *Neuropsychiatric Disease and Treatment*, pages 2905–2915.

Lin, H. (2022). Bayesian Epistemology. In Zalta, E. N. and Nodelman, U., editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edition.

Lin, S., Wang, Y., Zhang, L., Chu, Y., Liu, Y., Fang, Y., Jiang, M., Wang, Q., Zhao, B., Xiong, Y., et al. (2022). Mdf-sa-ddi: predicting drug–drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism. *Briefings in Bioinformatics*, 23(1):bbab421.

Lin, X., Quan, Z., Wang, Z.-J., Ma, T., and Zeng, X. (2020). Kgnn: Knowledge graph neural network for drug-drug interaction prediction. In *IJCAI*, volume 380, pages 2739–2745.

Liu, L., Jiang, H., He, P., Chen, W., Liu, X., Gao, J., and Han, J. (2019). On the variance of the adaptive learning rate and beyond. In *International Conference on Learning Representations*.

Lobdell, D. T., Jagai, J. S., Rappazzo, K., and Messer, L. C. (2011). Data sources for an environmental quality index: availability, quality, and utility. *American journal of public health*, 101(S1):S277–S285.

Long, Y., Chen, Q., Larsson, H., and Rzhetsky, A. (2021). Observable variations in human sex ratio at birth. *PLoS computational biology*, 17(12):e1009586.

Long, Y., Khan, A., and Rzhetsky, A. (2023). Peri-and post-natal risk factors associated with health of newborns. *medRxiv*.

Long, Y., Pan, H., Zhang, C., Song, H. T., Kondor, R., and Rzhetsky, A. (2022). Molecular fingerprints are a simple yet effective solution to the drug–drug interaction problem. *ICML Workshop on Computational Biology*.

Loshchilov, I. and Hutter, F. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

Ludvigsson, J. F., Andersson, E., Ekbom, A., Feychting, M., Kim, J.-L., Reuterwall, C., Heurgren, M., and Olausson, P. O. (2011). External review and validation of the swedish national inpatient register. *BMC public health*, 11(1):450.

Lykins, J., Wang, K., Wheeler, K., Clouser, F., Dixon, A., El Bissati, K., Zhou, Y., Lyttle, C., Rzhetsky, A., and McLeod, R. (2016). Understanding toxoplasmosis in the united states through "large data" analyses. *Reviews of Infectious Diseases*, 63(4):468–475.

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

Maenner, M. J., Shaw, K. A., Baio, J., et al. (2020). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2016. *MMWR Surveillance summaries*, 69(4):1.

Mancini, A. D., Littleton, H. L., and Grills, A. E. (2016). Can people benefit from acute stress? social support, psychological improvement, and resilience after the Virginia Tech campus shootings. *Clinical Psychological Science*, 4(3):401–417.

Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge University Press.

McCain, K. and Poston, T. (2017). *Best explanations: new essays on inference to the best explanation*. Oxford University Press.

McDonnell, L., Gilkes, A., Ashworth, M., Rowland, V., Harries, T. H., Armstrong, D., and White, P. (2021). Association between antibiotics and gut microbiome dysbiosis in children: systematic review and meta-analysis. *Gut Microbes*, 13(1):1870402.

Meincke, A. S. (2022). One or two? a process view of pregnancy. *Philosophical Studies*, 179(5):1495–1521.

Meincke, A. S. and Dupré, J. (2020). *Biological Identity: Perspectives from Metaphysics and the Philosophy of Biology*. Routledge.

Melamed, R. D. and Rzhetsky, A. (2018). Patchwork of contrasting medication cultures across the USA. *Nature Communications*, 9(1):4022.

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

Melville, J. M. and Moss, T. J. (2013). The immune consequences of preterm birth. *Frontiers in neuroscience*, 7:79.

Messer, L. C., Jagai, J. S., Rappazzo, K. M., and Lobdell, D. T. (2014). Construction of an environmental quality index for public health research. *Environmental Health*, 13(1):39.

Messinger, C. J., Lipsitch, M., Bateman, B. T., He, M., Huybrechts, K. F., MacDonald, S., Mogun, H., Mott, K., and Hernández-Díaz, S. (2020). Association between congenital cytomegalovirus and the prevalence at birth of microcephaly in the united states. *JAMA pediatrics*, 174(12):1159–1167.

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., and Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS one*, 9(7):e103537.

Morais, L. H., Schreiber, H. L., and Mazmanian, S. K. (2021). The gut microbiota–brain axis in behaviour and brain disorders. *Nature Reviews Microbiology*, 19(4):241–255.

Morgan, H. L. (1965). The generation of a unique machine description for chemical structures-a technique developed at chemical abstracts service. *Journal of chemical documentation*, 5(2):107–113.

Morimoto, J. (2020). Bateman (1948): was it all wrong? a comment on hoquet (2020). *Animal Behaviour*, 168:e1–e4.

Muenchhoff, M. and Goulder, P. J. (2014). Sex differences in pediatric infectious diseases. *The Journal of infectious diseases*, 209(suppl_3):S120–S126.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. The MIT Press.

Myung, I. J. and Pitt, M. A. (1997). Applying occam's razor in modeling cognition: A bayesian approach. *Psychonomic bulletin & review*, 4:79–95.

Nagpal, J. and Cryan, J. F. (2021). Microbiota-brain interactions: Moving toward mechanisms in model organisms. *Neuron*.

Narla, S. and Silverberg, J. I. (2019). Association between atopic dermatitis and autoimmune disorders in us adults and children: a cross-sectional study. *Journal of the American Academy of Dermatology*, 80(2):382–389.

Nemoto, N., Kalkbrenner, A., Friedman, L., and Jagai, J. (2021). Ambient air pollution (pm2. 5 & o3) in relation to adhd in nyc children age 3-13 years. In *ISEE Conference Abstracts*, volume 2021.

Neuman, H., Forsythe, P., Uzan, A., Avni, O., and Koren, O. (2018). Antibiotics in early life: dysbiosis and the damage done. *FEMS microbiology reviews*, 42(4):489–499.

Nicholson, D. J. and Dupré, J. (2018). *Everything flows: towards a processual philosophy of biology*. Oxford University Press.

Nieminen, P., Lehtiniemi, H., Huusko, A., Vähäkangas, K., and Rautio, A. (2013). Polychlorinated biphenyls (pcbs) in relation to secondary sex ratio–a systematic review of published studies. *Chemosphere*, 91(2):131–138.

Nyamabo, A. K., Yu, H., Liu, Z., and Shi, J.-Y. (2022). Drug–drug interaction prediction with learnable size-adaptive molecular substructures. *Briefings in Bioinformatics*, 23(1):bbab441.

Nyamabo, A. K., Yu, H., and Shi, J.-Y. (2021). Ssi–ddi: substructure–substructure interactions for drug–drug interaction prediction. *Briefings in Bioinformatics*, 22(6):bbab133.

Ojeda, J., Ávila, A., and Vidal, P. M. (2021). Gut microbiota interaction with the central nervous system throughout life. *Journal of Clinical Medicine*, 10(6):1299.

Olin, A., Henckel, E., Chen, Y., Lakshmikanth, T., Pou, C., Mikes, J., Gustafsson, A., Bernhardsson, A. K., Zhang, C., Bohlin, K., et al. (2018). Stereotypic immune system development in newborn children. *Cell*, 174(5):1277–1292.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., and Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, 112(16):E2102–E2111.

Pang, S., Zhang, Y., Song, T., Zhang, X., Wang, X., and Rodriguez-Patón, A. (2022). Amde: a novel attention-mechanism-based multidimensional feature encoder for drug–drug interaction prediction. *Briefings in Bioinformatics*, 23(1):bbab545.

Parenti, I., Rabaneda, L. G., Schoen, H., and Novarino, G. (2020). Neurodevelopmental disorders: from genetics to functional pathways. *Trends in Neurosciences*, 43(8):608–621.

Parker, G. A. (2014). The sexual cascade and the rise of pre-ejaculatory (darwinian) sexual selection, sex roles, and sexual conflict. *Cold Spring Harbor perspectives in biology*, 6(10):a017509.

Parker, G. A. and Pizzari, T. (2015). Sexual selection: the logical imperative. In *Current perspectives on sexual selection*, pages 119–163. Springer.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc.

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

Pearl, J. (2009). *Causality*. Cambridge university press.

Perez-Muñoz, M. E., Arrieta, M.-C., Ramer-Tait, A. E., and Walter, J. (2017). A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. *microbiome*, 5(1):1–19.

Polanczyk, G., De Lima, M. S., Horta, B. L., Biederman, J., and Rohde, L. A. (2007). The worldwide prevalence of adhd: a systematic review and metaregression analysis. *American journal of psychiatry*, 164(6):942–948.

Postma, E., Heinrich, F., Koller, U., Sardell, R. J., Reid, J. M., Arcese, P., and Keller, L. F. (2011). Disentangling the effect of genes, the environment and chance on sex ratio variation in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences*, 278(1720):2996–3002.

Potochnik, A. (2017). *Idealization and the Aims of Science*. University of Chicago Press.

Potochnik, A. (2020). Idealization and many aims. *Philosophy of science*, 87(5):933–943.

Pradeu, T. (2020). *Philosophy of immunology*. Cambridge University Press.

Preston, B. T., Stevenson, I. R., Pemberton, J. M., and Wilson, K. (2001). Dominant rams lose out by sperm depletion. *Nature*, 409(6821):681–682.

Proal, A. D., Lindseth, I. A., and Marshall, T. G. (2017). Microbe-microbe and host-microbe interactions drive microbiome dysbiosis and inflammatory processes. *Discovery medicine*, 23(124):51–60.

Puertas, A., Magan-Fernandez, A., Blanc, V., Revelles, L., O'Valle, F., Pozo, E., León, R., and Mesa, F. (2018). Association of periodontitis with preterm birth and low birth weight: a comprehensive review. *The Journal of Maternal-Fetal & Neonatal Medicine*, 31(5):597–602.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Renz, H., Brandtzaeg, P., and Hornef, M. (2012). The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. *Nature Reviews Immunology*, 12(1):9–23.

Rogers, D. and Hahn, M. (2010). Extended-connectivity fingerprints. *Journal of chemical information and modeling*, 50(5):742–754.

Root, A., Brown, J. P., Forbes, H. J., Bhaskaran, K., Hayes, J., Smeeth, L., and Douglas, I. J. (2019). Association of relative age in the school year with diagnosis of intellectual disability, attention-deficit/hyperactivity disorder, and depression. *JAMA pediatrics*, 173(11):1068–1075.

Rose, D. R., Careaga, M., Van de Water, J., McAllister, K., Bauman, M. D., and Ashwood, P. (2017). Long-term altered immune responses following fetal priming in a non-human primate model of maternal immune activation. *Brain, behavior, and immunity*, 63:60–70.

Rosenberg, A. and McIntyre, L. (2020). *Philosophy of biology: A contemporary introduction, 4th edition*. Routledge.

Roubaud-Baudron, C., Ruiz, V. E., Swan Jr, A. M., Vallance, B. A., Ozkul, C., Pei, Z., Li, J., Battaglia, T. W., Perez-Perez, G. I., and Blaser, M. J. (2019). Long-term effects of early-life antibiotic exposure on resistance to subsequent bacterial infection. *MBio*, 10(6):e02820–19.

Roughgarden, J. (2015). Sexual selection: Is anything left? *Current Perspectives on Sexual Selection*, pages 85–102.

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. K. (2017). Bayesian computing with inla: a review. *Annual Review of Statistics and Its Application*, 4:395–421.

Ryu, J. Y., Kim, H. U., and Lee, S. Y. (2018). Deep learning improves prediction of drug–drug and drug–food interactions. *Proceedings of the National Academy of Sciences*, 115(18):E4304–E4311.

Sahu, S. K. and Anand, A. (2018). Drug-drug interaction extraction from biomedical texts using long short-term memory network. *Journal of biomedical informatics*, 86:15–24.

Säilynoja, T., Bürkner, P.-C., and Vehtari, A. (2022). Graphical test for discrete uniformity and its applications in goodness-of-fit evaluation and multiple sample comparison. *Statistics and Computing*, 32(2):1–21.

Sam, Q. H., Chang, M. W., and Chai, L. Y. A. (2017). The fungal mycobiome and its interaction with gut bacteria in the host. *International journal of molecular sciences*, 18(2):330.

Samartsidis, P., Seaman, S. R., Presanis, A. M., Hickman, M., and De Angelis, D. (2018). Review of methods for assessing the causal effect of binary interventions from aggregate time-series observational data. *arXiv preprint arXiv:1804.07683*.

Sariola, S. and Gilbert, S. F. (2020). Toward a symbiotic perspective on public health: recognizing the ambivalence of microbes in the anthropocene. *Microorganisms*, 8(5):746.

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

Scherb, H., Voigt, K., and Kusmierz, R. (2015). Ionizing radiation and the human gender proportion at birth—a concise review of the literature and complementary analyses of historical and recent data. *Early human development*, 91(12):841–850.

Schluter, J., Peled, J. U., Taylor, B. P., Markey, K. A., Smith, M., Taur, Y., Niehus, R., Staffas, A., Dai, A., Fontana, E., et al. (2020). The gut microbiota is associated with immune cell dynamics in humans. *Nature*, 588(7837):303–307.

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

Scott, S. L. (2019). *bsts: Bayesian Structural Time Series*. R package version 0.9.2.

Scott, S. L. and Varian, H. R. (2013). Predicting the present with bayesian structural time series. *Available at SSRN 2304426*.

Sevelsted, A., Stokholm, J., Bønnelykke, K., and Bisgaard, H. (2015). Cesarean section and chronic immune disorders. *Pediatrics*, 135(1):e92–e98.

Shao, J. (2003). *Mathematical statistics*. Springer Science & Business Media.

Sharon, G., Cruz, N. J., Kang, D.-W., Gandal, M. J., Wang, B., Kim, Y.-M., Zink, E. M., Casey, C. P., Taylor, B. C., Lane, C. J., et al. (2019). Human gut microbiota from autism spectrum disorder promote behavioral symptoms in mice. *Cell*, 177(6):1600–1618.

Shekhar, S. and Petersen, F. C. (2020). The dark side of antibiotics: Adverse effects on the infant immune defense against infection. *Frontiers in pediatrics*, page 651.

Sjölander, A. and Vansteelandt, S. (2019). Frequentist versus bayesian approaches to multiple testing. *European journal of epidemiology*, 34:809–821.

Slob, E. M., Brew, B. K., Vijverberg, S. J., Dijs, T., Van Beijsterveldt, C. E., Koppelman, G. H., Bartels, M., Dolan, C. V., Larsson, H., Lundström, S., et al. (2021). Early-life antibiotic use and risk of attention-deficit hyperactivity disorder and autism spectrum disorder: results of a discordant twin study. *International journal of epidemiology*, 50(2):475–484.

Sokol, H., Leducq, V., Aschard, H., Pham, H.-P., Jegou, S., Landman, C., Cohen, D., Liguori, G., Bourrier, A., Nion-Larmurier, I., et al. (2017). Fungal microbiota dysbiosis in ibd. *Gut*, 66(6):1039–1048.

Song, S. (2018). 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*, 39(3):336–342.

Sorboni, S. G., Moghaddam, H. S., Jafarzadeh-Esfehani, R., and Soleimanpour, S. (2022). A comprehensive review on the role of the gut microbiome in human neurological disorders. *Clinical Microbiology Reviews*, 35(1):e00338–20.

Sprenger, J. and Hartmann, S. (2019). *Bayesian philosophy of science*. oxford university press.

Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.2.

Stinson, L. F., Payne, M. S., and Keelan, J. A. (2018). A critical review of the bacterial baptism hypothesis and the impact of cesarean delivery on the infant microbiome. *Frontiers in medicine*, 5:135.

Stokholm, J., Thorsen, J., Blaser, M. J., Rasmussen, M. A., Hjelmsø, M., Shah, S., Christensen, E. D., Chawes, B. L., Bønnelykke, K., Brix, S., et al. (2020). Delivery mode and gut microbial changes correlate with an increased risk of childhood asthma. *Science translational medicine*, 12(569):eaax9929.

Straub, L., Bateman, B. T., Hernandez-Diaz, S., York, C., Zhu, Y., Suarez, E. A., Lester, B., Gonzalez, L., Hanson, R., Hildebrandt, C., et al. (2021). Validity of claims-based algorithms to identify neurodevelopmental disorders in children. *Pharmacoepidemiology and Drug Safety*, 30(12):1635–1642.

Su, X., Hu, L., You, Z., Hu, P., and Zhao, B. (2022). Attention-based knowledge graph representation learning for predicting drug-drug interactions. *Briefings in Bioinformatics*.

Suárez, J. and Stencel, A. (2020). A part-dependent account of biological individuality: Why holobionts are individuals and ecosystems simultaneously. *Biological Reviews*, 95(5):1308–1324.

Sucksdorff, M., Lehtonen, L., Chudal, R., Suominen, A., Gissler, M., and Sourander, A. (2018). Lower apgar scores and caesarean sections are related to attention-deficit/hyperactivity disorder. *Acta Paediatrica*, 107(10):1750–1758.

Sultan, S. E. (2021). Eco-evo-devo. In Nuño de la Rosa, L. and Müller, G. B., editors, *Evolutionary Developmental Biology: A Reference Guide*, pages 1165–1177. Springer International Publishing, Cham.

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

Takeshita, C. (2022). From mother/fetus to holobiont (s): A material feminist ontology of the pregnant body. *Catalyst: Feminism, Theory, Technoscience*, 3(1).

Tamer, G. S., Dundar, D., Yalug, I., Caliskan, S., Yazar, S., and Aker, A. (2008). The schizophrenia and toxoplasma gondii connection: infectious, immune or both? *Advances in therapy*, 25:703–709.

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

Terrell, M. L., Hartnett, K. P., and Marcus, M. (2011). Can environmental or occupational hazards alter the sex ratio at birth? a systematic review. *Emerging health threats journal*, 4(1):7109.

Thion, M. S., Low, D., Silvin, A., Chen, J., Grisel, P., Schulte-Schrepping, J., Blecher, R., Ulas, T., Squarzoni, P., Hoeffel, G., et al. (2018). Microbiome influences prenatal and adult microglia in a sex-specific manner. *Cell*, 172(3):500–516.

Tick, B., Bolton, P., Happé, F., Rutter, M., and Rijsdijk, F. (2016). Heritability of autism spectrum disorders: a meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry*, 57(5):585–595.

Trivers, R. (1972). Parental investment and sexual selection. *Sexual Selection & the Descent of Man, Aldine de Gruyter, New York*, pages 136–179.

Trivers, R. L. and Willard, D. E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science*, 179(4068):90–92.

Van de Guchte, M., Blottière, H. M., and Doré, J. (2018). Humans as holobionts: implications for prevention and therapy. *Microbiome*, 6(1):1–6.

van Dongen, N., Sprenger, J., and Wagenmakers, E.-J. (2023). A bayesian perspective on severity: Risky predictions and specific hypotheses. *Psychonomic Bulletin & Review*, 30(2):516–533.

Vandenbroucke, J. P., Elm, E. v., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., Egger, M., and Initiative, S. (2007). Strengthening the reporting of observational studies in epidemiology (strobe): explanation and elaboration. *Annals of internal medicine*, 147(8):W–163.

Vangay, P., Ward, T., Gerber, J. S., and Knights, D. (2015). Antibiotics, pediatric dysbiosis, and disease. *Cell host & microbe*, 17(5):553–564.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., and Gelman, A. (2022a). loo: Efficient leave-one-out cross-validation and waic for bayesian models. R package version 2.5.0.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., and Gelman, A. (2022b). loo: Efficient leave-one-out cross-validation and waic for bayesian models. R package version 2.5.0.

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

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

Vehtari, A. and Lampinen, J. (2002). Bayesian model assessment and comparison using cross-validation predictive densities. *Neural computation*, 14(10):2439–2468.

Vehtari, A., Simpson, D., Gelman, A., Yao, Y., and Gabry, J. (2015). Pareto smoothed importance sampling. *arXiv preprint arXiv:1507.02646*.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. (2018). Graph Attention Networks. *International Conference on Learning Representations*. accepted as poster.

Vilar, S., Harpaz, R., Uriarte, E., Santana, L., Rabadan, R., and Friedman, C. (2012). Drug—drug interaction through molecular structure similarity analysis. *Journal of the American Medical Informatics Association*, 19(6):1066–1074.

Vilar, S., Uriarte, E., Santana, L., Lorberbaum, T., Hripcsak, G., Friedman, C., and Tatonetti, N. P. (2014). Similarity-based modeling in large-scale prediction of drug-drug interactions. *Nature protocols*, 9(9):2147–2163.

Wade, M. J. and Shuster, S. M. (2005). Don't throw bateman out with the bathwater! *Integrative and Comparative Biology*, 45(5):945–951.

Walker, W. A. (2017). The importance of appropriate initial bacterial colonization of the intestine in newborn, child, and adult health. *Pediatric research*, 82(3):387–395.

Wang, J., Liu, X., Shen, S., Deng, L., and Liu, H. (2022). Deepdds: deep graph neural network with attention mechanism to predict synergistic drug combinations. *Briefings in Bioinformatics*, 23(1):bbab390.

West, S. (2009). Sex allocation. In *Sex Allocation*. Princeton University Press.

Winkelmann, R. (2008). *Econometric analysis of count data, 5th edition*. Springer Science & Business Media.

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic acids research*, 46(D1):D1074–D1082.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.

World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization.

Wu, X., Xia, Y., He, F., Zhu, C., and Ren, W. (2021). Intestinal mycobiota in health and diseases: from a disrupted equilibrium to clinical opportunities. *Microbiome*, 9(1):1–18.

Xiong, X., Harville, E. W., Buekens, P., Mattison, D. R., Elkind-Hirsch, K., and Pridjian, G. (2008). Exposure to Hurricane Katrina, post-traumatic stress disorder and birth outcomes. *The American journal of the medical sciences*, 336(2):111–115.

Xu, H., Liu, M., Cao, J., Li, X., Fan, D., Xia, Y., Lu, X., Li, J., Ju, D., and Zhao, H. (2019a). The dynamic interplay between the gut microbiota and autoimmune diseases. *Journal of immunology research*, 2019.

Xu, L.-l., Zhang, X., Zhou, G.-l., Jiang, C.-m., Jiang, H.-y., and Zhou, Y.-y. (2020). Meta-analysis found that studies may have overestimated caesarean section risks for attention-deficit hyperactivity disorder by ignoring confounding factors. *Acta Paediatrica*, 109(2):258–265.

Xu, N., Wang, P., Chen, L., Tao, J., and Zhao, J. (2019b). Mr-gnn: Multi-resolution and dual graph neural network for predicting structured entity interactions. *arXiv preprint arXiv:1905.09558*.

Yacoub, R., Jacob, A., Wlaschin, J., McGregor, M., Quigg, R. J., and Alexander, J. J. (2018). Lupus: the microbiome angle. *Immunobiology*, 223(6-7):460–465.

Yau, S.-y., Lu, H.-x., Lee, T. M., Guo, H., Chan, C. C., et al. (2021). Pm2. 5 as a potential risk factor for autism spectrum disorder: its possible link to neuroinflammation, oxidative stress and changes in gene expression. *Neuroscience & Biobehavioral Reviews*, 128:534–548.

Yin, Q., Cao, X., Fan, R., Liu, Q., Jiang, R., and Zeng, W. (2022). Deepdrug: A general graph-based deep learning framework for drug-drug interactions and drug-target interactions prediction. *bioRxiv*.

Yu, H.-y., Zhou, Y.-y., Pan, L.-y., Zhang, X., and Jiang, H.-y. (2022a). Early life antibiotic exposure and the subsequent risk of autism spectrum disorder and attention deficit hyperactivity disorder: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, 52(5):2236–2246.

Yu, L. W., Agirman, G., and Hsiao, E. Y. (2022b). The gut microbiome as a regulator of the neuroimmune landscape. *Annual Review of Immunology*, 40:143–167.

Zagidullin, B., Wang, Z., Guan, Y., Pitkänen, E., and Tang, J. (2021). Comparative analysis of molecular fingerprints in prediction of drug combination effects. *Briefings in bioinformatics*, 22(6):bbab291.

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., and Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, 15(5):778–790.

Zhang, H., Dahlén, T., Khan, A., Edgren, G., and Rzhetsky, A. (2020a). Measurable health effects associated with the daylight saving time shift. *PLoS computational biology*, 16(6):e1007927.

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

Zhang, T., Brander, G., Mantel, Ä., Kuja-Halkola, R., Stephansson, O., Chang, Z., Larsson, H., Mataix-Cols, D., and de la Cruz, L. F. (2021). Assessment of cesarean delivery and neurodevelopmental and psychiatric disorders in the children of a population-based swedish birth cohort. *JAMA network open*, 4(3):e210837–e210837.

Zhang, T., Leng, J., and Liu, Y. (2020b). Deep learning for drug–drug interaction extraction from the literature: a review. *Briefings in bioinformatics*, 21(5):1609–1627.

Zhang, T., Sidorchuk, A., Sevilla-Cermeño, L., Vilaplana-Pérez, A., Chang, Z., Larsson, H., Mataix-Cols, D., and de la Cruz, L. F. (2019b). Association of cesarean delivery with risk of neurodevelopmental and psychiatric disorders in the offspring: a systematic review and meta-analysis. *JAMA network open*, 2(8):e1910236–e1910236.

Zhang, X., Zou, Y., and Shi, W. (2017). Dilated convolution neural network with leakyrelu for environmental sound classification. In *2017 22nd international conference on digital signal processing (DSP)*, pages 1–5. IEEE.

Zhao, Z., Yang, Z., Luo, L., Lin, H., and Wang, J. (2016). Drug drug interaction extraction from biomedical literature using syntax convolutional neural network. *Bioinformatics*, 32(22):3444–3453.

Zheng, D., Liwinski, T., and Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell research*, 30(6):492–506.

Zheng, W., Lin, H., Luo, L., Zhao, Z., Li, Z., Zhang, Y., Yang, Z., and Wang, J. (2017). An attention-based effective neural model for drug-drug interactions extraction. *BMC bioinformatics*, 18(1):1–11.

Zietsch, B. P., Walum, H., Lichtenstein, P., Verweij, K. J., and Kuja-Halkola, R. (2020). 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*, 287(1921):20192849.

Zietsch, B. P., Walum, H., Lichtenstein, P., Verweij, K. J., and Kuja-Halkola, R. (2021). When theory cannot explain data, the theory needs rethinking. invited replies to: Orzack SH, Hardy ICW. 2021, and Lehtonen J. 2021. *Proceedings of the Royal Society B*, 288(1947):20210304.

Zitnik, M., Agrawal, M., and Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*, 34(13):i457–i466.