Exploring spatio-temporal patterns in emergency department use for mental health reasons from children and adolescents in Alberta, Canada

by

Michelle Louise Thiessen

B.Sc., University of Manitoba, 2016

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Statistics and Actuarial Science
Faculty of Science

© Michelle Louise Thiessen 2018
SIMON FRASER UNIVERSITY
Summer 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

Approval

Name:	**Michelle Louise Thiessen**
Degree:	**Master of Science (Statistics)**
Title:	**Exploring spatio-temporal patterns in emergency department use for mental health reasons from children and adolescents in Alberta, Canada**
Examining Committee:	**Chair:** Jinko Graham Professor **X. Joan Hu** Senior Supervisor Professor **Rhonda J. Rosychuk** Supervisor Professor Division of Infectious Diseases Department of Pediatrics University of Alberta **Brad McNeney** Internal Examiner Associate Professor
Date Defended:	**July 26, 2018**

Abstract

This project analyses mental health related emergency department visits from children and adolescents in Alberta, Canada to understand the spatio-temporal patterns and identify risk factors. The data are extracted for the period 2002-2011 from the provincial health administrative data systems of Alberta. A descriptive data analysis is presented and then generalized linear models are explored to model the spatio-temporal pattern of the emergency department visit counts. The seasonal effect is examined using seasonal factors, sine and cosine functions and cyclic cubic smoothing splines. The spatial and temporal correlation structures are modelled using autoregressive model of order 1 and conditionally autoregressive model random effects. Demographic risk factors and their association with the frequency of mental health related emergency department visits is examined. Estimates of the model parameters are obtained and model diagnostics are performed to assess the fit of the model. Age, gender and proxy for socio-economic status are found to be important risk factors. The proposed model can be used as a predictive model to help identify regions and groups at a higher risk for mental health related emergency department visits.

Keywords: generalized linear model; mixed effects; regression analysis; seasonal effect

Acknowledgements

First and foremost I would like to thank my senior supervisor Professor Joan Hu. Thank you Joan for your patience, encouragement and support over the past two years. Your guidance has been instrumental to my success. I want to thank my supervisor Professor Rhonda Rosychuk for giving me the opportunity to work with this data and for all her help and patience with running my code. I want to extend a special thanks to Professor Brad McNeney for taking the time to serve on my committee.

I would like to thank all of my professors here at Simon Fraser University for their exceptional lectures. Thank you to all my fellow graduate students, particularly Sarah Bailey, JinCheol Choi, Richard Hsia, Grace Hsu, Jinwan Kim, Lillian Lin, Yue Ma, Payman Nickchi, Will Ruth, Trevor Thomson and Ran Wang. I want to thank all of you for your support and thought-provoking conversations had over tea. My time here would not have been the same without all of you!

Last but not least, I want to thank my family. To my parents, I wouldn't be where I am today without your love and constant words of encouragement. To my siblings, thank you for always being there when I need you the most. I couldn't ask for better siblings.

Disclaimer

This project is based in part on data provided by Alberta Health. The interpretation and conclusions contained herein are those of the researchers and do not necessarily represent the views of the Government of Alberta. Neither the government nor Alberta Health expresses any opinion in relation to this project.

Table of Contents

Approval ii

Abstract iii

Acknowledgements iv

Disclaimer v

Table of Contents vi

List of Tables viii

List of Figures ix

1 Introduction 1

1.1 Background 1

1.2 Objectives 2

1.3 Outline 3

2 Alberta Pediatric Mental Health Care (PMHC) Dataset 4

2.1 Overview of PMHC Dataset 4

2.2 Descriptive Analysis 5

2.3 Spatial Autocorrelation 8

2.4 Temporal Autocorrelation 10

3 Generalized Linear Mixed Models (GLMM) 11

3.1 Notation 11

3.2 Modelling 12

3.2.1 Modelling temporal patterns 13

3.2.2 Modelling spatial-temporal patterns 14

4 Regression Analysis Under GLMM 16

4.1 Overall Analysis Procedure 16

4.2 Temporal Analysis 17

4.2.1 Risk factors/exposures effects 17
4.2.2 Estimates of average counts 18
4.2.3 Seasonality and time trend 19
4.2.4 Model checking 20
4.2.5 Analysis of residuals 20
4.2.6 Proposed model given temporal analysis 21
4.3 Spatio-Temporal Analysis 21
4.3.1 Motivation for Integrated Nested Laplace Approximation (INLA) . . 23
4.3.2 Model exploration 24
4.3.3 Proposed model given spatio-temporal analysis 27

5 Final Remarks **29**
5.1 Summary 29
5.2 Future Investigation 30

Bibliography **31**

Appendix A List of Tables **33**

Appendix B List of Figures **40**

List of Tables

Table A.1 Summary of the MHED counts and subjects with an MHED visit counts by demographic information. 33
Table A.2 Summary of the MHED counts and subjects with an MHED visit counts by RHA. 34
Table A.3 Weighted average percentages of children/youth with an MHED visit in Alberta 35
Table A.4 Summary of the goodness of variance fit and tabular accuracy index measures for different generalizing techniques. 35
Table A.5 Global Moran's I and EBI statistics and their p-values for number of MHED visits. 35
Table A.6 Data format of aggregation at the sRHA-level. 36
Table A.7 Summary of dispersion parameter estimate for the three proposed temporal models and three coefficient estimation settings. 36
Table A.8 Summary of the constant coefficient estimates for the exposure and risk factors from the seasonal factors, trigonometric functions and smoothing splines models. 36
Table A.9 Summary of the coefficient estimates for model ST1 and model ST2 with a trigonometric seasonal effect. 37
Table A.10 AICs and BICs for three different settings of model ST3, a nested generalized linear mixed model. 37
Table A.11 WAIC and DICs for model ST4 with highly vague priors. 37
Table A.12 WAIC and DICs for model ST4 with weakly informative priors. . . . 38
Table A.13 Summary of model ST4 with scaled t and AR(1), sRHA, pSES, seasonal and time trend random effects. 39

List of Figures

Figure B.1 Plots of MHED visit counts and subject visit counts by RHA over time (yearly). 40
Figure B.2 Plots of MHED visits per 1000 RHA population over months (top), days (middle), and hour of day (bottom) aggregated over all fiscal years. 41
Figure B.3 Choropleth maps of number of MHED visits per 1000 sRHA population over 9 fiscal years. 42
Figure B.3 Choropleth maps of number of ED visits per 1000 sRHA population over 9 fiscal years. 43
Figure B.3 Choropleth maps of number of MHED visits per 1000 sRHA population over 9 fiscal years. 44
Figure B.4 ACF plots of monthly MHED visits over nine fiscal years for each RHA. 45
Figure B.5 PACF plots of monthly MHED visits over nine fiscal years for each RHA. 46
Figure B.6 Distribution of the counts for the number of MHED visits and population counts. 47
Figure B.7 Part I of the plots of the coefficient estimates for the seasonal factors, trigonometric functions and smoothing splines models. 48
Figure B.8 Part II of the plots of the coefficient estimates for the seasonal factors, trigonometric functions and smoothing splines models. 49
Figure B.9 Plots of the estimated average MHED counts per 1000 sRHA population for select sRHAs for combinations of females, teenagers and the three pSES categories for the three estimation settings. 50
Figure B.10 Plots of the estimated average MHED counts per 1000 sRHA population for select sRHAs for combinations of females, teenagers and the three pSES categories for the three proposed temporal models. 51
Figure B.11 Plot of the seasonal coefficients for the seasonal factors model. . . 52
Figure B.12 Plots of the seasonal coefficients and time coefficient over the time unit (28 days) for the seasonal factors and trigonometric models. . 53

Figure B.13 Plots of the smoothing splines for the blocks and time by RHA from model T3. 54
Figure B.14 Deviance residuals vs fitted values of the trigonometric functions temporal model under the three estimation settings. 55
Figure B.15 Deviance residuals vs fitted values of the generalized linear mixed model with nested random effects. 56
Figure B.16 CPO and PIT values for Model ST4: AR(1), CAR, pSES and seasonal and time trend random effects. 57
Figure B.17 Plots of true and fitted MHED visits per 1000 sRHA population for females aged greater than 12 in sRHA R608 who are regular plan participants (RPP), treaty status or income supported. 58
Figure B.18 Plots of true and fitted MHED visits per 1000 sRHA population for females aged greater than 12 in sRHA R804 who are regular plan participants (RPP), treaty status or income supported. 59

Chapter 1

Introduction

1.1 Background

Mental health is a relatively understudied subject matter, yet it is estimated that 10-20% of Canadian children and youth are affected by a mental illness or disorder. Unfortunately, only 20% of those who need mental health services currently receive them (Leitch, 2007). In an annual report by the Alberta Health Services for the years 2016-2017, it is reported that there are well over 2 million visits paid to Alberta emergency departments every year. In Alberta, there is a lack of community-based supports for mental health services, and inpatient and outpatient treatment options. This often leads to families seeking help in emergency departments (ED) during crises (Newton et al., 2011). Mental health related emergency department (MHED) visits can be avoidable if treatment is sought out before the crisis point is reached. Learning more about where EDs are experiencing high amounts of MHED visits and their trend can help the province of Alberta understand where more resources are needed and reduce overcrowding in EDs.

This project uses the pediatric mental health care (PMHC) dataset that was extracted from four population-based administrative databases in Alberta. Previous work using the PMHC dataset has found evidence that the MHED visits for subsets of the Alberta population cluster over space and time. For each of the following citations, a cluster was defined as a geographic area and time period that had a statistically higher number of cases of a disease than expected. Mariathas and Rosychuk (2015) studied three different spatial cluster detection methods that differed in their choice of distributional assumption. ED presentations by children and youth aged less than 18 years old for substance use during April 1, 2007 to March 31, 2008 were used. Statistically significant clusters were found in northern Alberta, parts of the Edmonton region, and southwestern Alberta for all three detection methods. Rosychuk et al. (2015) used ED presentations due to a mood disorder for Albertans aged 10-17. The Kulldorff-Nagarwalla (KN) spatial scan test (Kulldorff and Nagarwalla, 1995) found three potential clusters over space and time in the majority of northern Alberta between 2007 and 2011, in a single sub-regional health authority (sRHA)

in the southwest part of the Central Zone between 2005 and 2009, and another in the Central Zone between 2008 and 2011. Newton et al. (2016) identified clusters over space and time for Alberta adolescents aged 15-17 years during 2002-2011 with an ED presentation for a mental or behavioral disorder secondary to alcohol or drug use. Using the KN spatial scan test, they found a cluster in the North, Edmonton and northwest Central zones between 2004 and 2008, in the western South and southern Calgary zones between 2007 to 2011, and in the northern South zone between 2006 and 2007. Rosychuk et al. (2016) used the KN spatial scan on Albertans aged 15-17 years during 2002-2011 who presented to an ED for self-harm and had no physician follow-up visit within 14 days post-ED visit. They found a cluster in northern Alberta from 2002-2006 and in southern Alberta between 2003-2007. These results motivate this project's objective of exploring the spatio-temporal pattern of the MHED data and to explore what impact demographic information has on the data.

There are some limitations with the PMHC dataset. Firstly, the records of the MHED visits are from individuals age 0 to 17 at the time of their visit between the dates April 1, 2002 and March 31, 2011. This is an issue since the data is both left and right truncated. We do not have information on whether the subjects in the data have had multiple visits before April 1, 2002, and whether they continued to have visits after March 31, 2011. We are restricted to the April 1, 2002 to March 31, 2011 observation window. Secondly, due to Alberta Health's privacy protocol, individual's birth dates are prevented from being released. Therefore we have incomplete information for the subjects' ages. Thirdly, the spatial information regarding the subject's place of residence and the emergency department visited is only available at the sub-Regional Health Authority (sRHA) level. The spatial information is not available as point data, but rather polygon data, so there is a loss of information. Lastly, the PMHC dataset does not have information regarding what location the subject came from to reach the ED. The point of origin may not necessarily be the subject's place of residence.

1.2 Objectives

In this project, the target population for statistical inference is the population of Alberta aged 0-18 at the fiscal year end. The sRHA population sizes, recorded at the fiscal year end, are incorporated in the modelling to account for the differences in population sizes between regions. A fiscal year for this data starts on April 1st and ends at the following March 31st. This project's goal is to see whether there is any significant spatio-temporal effect or pattern among the sub-Regional Health Authorities (sRHAs) of Alberta's total number of mental health related emergency department visits. In addition, we explore whether certain demographic groups (gender, age, and socio-economic status) are more susceptible to having MHED visits and whether this susceptibility varies over time and space.

1.3 Outline

The organization of the project is as follows. Chapter 2 starts by introducing the dataset and showcasing some descriptive analysis. The motivation for studying the spatio-temporal pattern of the MHED visits made by Albertans aged less than 18 is presented with descriptive plots and spatial and temporal autocorrelation findings. Chapter 3 introduces the notation, formulation, and proposed methodology. To capture the seasonality and trend of the data, a seasonal factors model, a trigonometric model and a generalized additive model are proposed. To observe the differences between regions and over time, the coefficients are allowed to vary over time and space. Lastly, to capture any remaining temporal and spatial correlation, random effects are proposed. Chapter 4 presents the results from the analysis using the proposed methodology. A generalized linear mixed model is the proposed final model given the analysis results. Final remarks and future work is given in Chapter 5.

Chapter 2

Alberta Pediatric Mental Health Care (PMHC) Dataset

Mental health is an issue that affects 1 in 5 Canadian youth and children (Leitch, 2007). The analysis of the Alberta Pediatric Mental Health Care (PMHC) dataset is motivated by this statistic. Section 2.1 begins by describing the PMHC dataset and the region descriptors used in the dataset. Section 2.2 conducts a descriptive analysis of the important risk factors associated with MHED visits and looks at how the MHED visit rates vary over time and space. Tests for spatial autocorrelation are examined and applied to the dataset in Section 2.3. Section 2.4 concludes by studying the temporal autocorrelation and seasonal pattern of the MHED visits.

2.1 Overview of PMHC Dataset

The pediatric mental health care (PMHC) dataset was taken from four population-based administrative databases in Alberta. The Ambulatory Care Classification System (ACCS), the Population Registry File (PRF), the Physicians Claims File (PCF), and the Hospitalizations Discharge Database (HDD). In this project, we focus on the data taken from the ACCS and PRF databases.

The individuals of interest for this project are Alberta residents who had at least one MHED visit during the observation window, April 1, 2002 to March 31, 2011. They are individuals who were younger than 18 years of age at the time of their MHED visit. Population data is used to account for the differences in population size between regions. The population-level individual is an Alberta resident aged 0 to 18 at the fiscal year end (March 31, 20XX). An Alberta resident is defined as an individual who is registered in the Alberta Health Care Insurance Plan (AHCIP).

The PMHC dataset has the following demographic information collected on the individuals. The individual's age at the time of the MHED visit and at fiscal year end, their proxy for socio-economic status (pSES), and their sex. The information collected on the

individual's place of residence is their residential region at the regional health authority (RHA) level and sub-regional health authority (sRHA) level, and the first three digits of their place of residence's postal code (FSA). The residential region categorization of being rural or urban is determined using the first two digits of the individuals postal code. This project focuses on data at the RHA and sRHA level and not the FSA level in order to match the population data format. Hospitalization information collected on the individuals is their diagnosis, triage level, triage date, discharge disposition, time and date of the MHED visit, amount of time to triage, amount of time from triage to physician assessment, total amount of time spent in ED, and the date of follow-up visits to physicians after the hospitalization. In addition, the sRHA of the ED the patient visited is known, but not the actual ED visited. There are 70 sRHAs and approximately 100 EDs in Alberta (Newton et al., 2016).

Individuals in this data can have multiple MHED visits which makes the data recurrent event data. In other words, the MHED visits made by the same individual are correlated. Newton et al. (2016) found that from April 1, 2002 to March 31, 2008, most children and youth (75.2%) had only one mental health related visit during the six-year period, while 24.8% of children and youth had multiple mental health-realted visits.

Alberta is divided into nine RHAs, which was further divided into 70 sRHAs. The motivation for the RHA boundaries was based on the delivery of health services in the region. Due to the RHAs varying dramatically in area and population size, comparisons between RHAs were difficult to make. This motivated the province of Alberta to divide the province into smaller regions. Postal codes were not ideal since the population size varied too much and many postal codes had very small population sizes which caused concerns over data confidentiality. Latitude and longitude blocks were an attractive option, however they would not mesh with the RHA boundaries. This motivated the use of sRHAs where the province could target certain regions within RHAs for specific programs, especially since some RHAs had already begun creating their own sub-regions. The criteria for the sRHAs was that they had to have a minimum total population size of 20,000. The final sRHAs were based on municipal boundaries in some RHAs and in others they were based on hospital catchment areas. The northern region of Alberta, RHA 9, is the only RHA whose sRHAs did not satisfy the minimum population size. This was due to the province not wanting the RHA to be limited to only 2 sRHAs, Fort McMurray and the rest of the RHA. The formation and discussion of the sRHAs began in 2003. The discussion summarized above can be found in the report: Calculating Small Area Analysis: Definition of Sub-region Geographic Units in Alberta, 2003.

2.2 Descriptive Analysis

A fiscal year is a one-year period starting from April 1, until the following year's March 31. For example, April 1, 2002, until March 31, 2003, is the first fiscal year of the PMHC

dataset. The age used in the following tables, plots, and analysis is the age at the fiscal year end. This means that some individuals who were aged 17 at the time of their MHED visits might have turned 18 by the end of the fiscal year. This unit of time for age is chosen to match the population data which also records age at fiscal year end.

Tables A.1 and A.2 summarize the demographic information and RHA counts by number of MHED visits and subjects, respectively, with data taken from the PRF and ACCS databases. From Table A.1, it can be seen that females, teenagers aged 13-18, and the proxy for socio-economic status (pSES) regular plan participant (RPP) comprise of the majority of MHED visits for all of the 9 fiscal years. Due to the low counts of MHED visits for individuals aged less than 13, the age groups 0-5 and 6-12 years old may have to be combined. The pSES categories government sponsored and welfare are similar groups by definition. These categories could be combined into a category referred to by income supported to account for the small number of MHED visits in the welfare group. Since the RHA regions R3 and R6 are the major urban areas of the Calgary and Edmonton area, respectively, they have the largest amount of MHED visits compared to the other regions, as seen in Table A.2. The proportion of repeat MHED visits is relatively small but has some variation over time and across the RHAs. The northern regions of Alberta, R8 and R9, tend to have a higher proportion of repeat MHED visits.

In order to get the average percentage of children and youth that have MHED visits per RHA, a weighted average is used because the RHA specific percentages are based on different population sizes. Let n_i denote the number of individuals with an MHED visit in RHA i and let w_i denote the population size of RHA i, $i = 1, ..., 9$. Let x_i denote the proportion of individuals in RHA i with an MHED visit, in other words $x_i = \frac{n_i}{w_i}$. The formula for the average rate of MHED visits is $\bar{x} = \sum_{i=1}^{9} \frac{x_i}{9}$. The formula for the weighted average rate is

$$\bar{x}_w = \frac{\sum_{i=1}^{9} w_i x_i}{\sum_{i=1}^{9} w_i} = \frac{\sum_{i=1}^{9} n_i}{\sum_{i=1}^{9} w_i}. \tag{2.1}$$

For the weighted average rates divided by demographic information, such as gender, the formula becomes $\bar{x}_{(j)w} = \frac{\sum_{i=1}^{9} n_{i(j)}}{\sum_{i=1}^{9} w_i}$, where j is the index of the demographic variable, e.g. $j =$Male.

The weighted averages for the overall MHED visits and separated by demographic groups is shown in Table A.3. The weighted average was found to be smaller than the average due to RHAs with smaller populations having higher rates of MHED visits and RHAs with larger populations having lower rates. The increase in the weighted average percentage of individuals with an MHED visit in Alberta in fiscal year 04/05, as seen in Table A.3, is most likely due to rural RHAs such as R7 which have an increasing trend in the rate of subjects with MHED visits. This is seen in Figure B.1. Looking at the demographic weighted averages in Table A.3, females and youth aged 13-18 have the largest percentage

of subjects with MHED visits when their RHA population size is taken into account. The pSES category RPP, regular plan participant, have the highest proportion of MHED visits. Those in the welfare category have the smallest proportion. Interestingly, the percentage of MHED visits in the pSES category government sponsored decreases from 0.10% to 0.05% in fiscal year 2009/2010. This is due to the number of Albertans aged 0-18 in the pSES category government sponsored decreasing from 113,628 in 2008/2009 to 23,263 in 2010/2011.

Figure B.1 shows the trend over yearly time of the number of MHED visits per 1000 RHA population and the trend of the number of subjects with an MHED visit per 1000 RHA population. Both rates follow similar trends over time. RHA zone R7 has an increasing trend whereas zones such as R3, the zone representing the Calgary region, show a fairly constant rate over time. The zones with the largest populations, R3 and R6, which contain Alberta's major urban areas, Calgary and Edmonton, have some of the lowest proportion of MHED visits which could be due to accessibility to options other than EDs or less severe mental illness.

Figure B.2 shows at what times the most and least amount ED visits occur and how they vary for the different RHAs. For monthly time, the RHAs appear to follow a similar trend. They all have a decrease in MHED visits in the summer months, July and August, and increases in the winter months, October and November, as well as March and May. This motivates exploring a cyclic seasonal effect for the MHED visits. For daily time, the RHA trends appear to vary, with some seeing increases in MHED visits on Saturday and others seeing decreases on Saturday. For time of day, all the RHAs see a decrease in MHED visits in the early hours of the morning, 1am to 6am, and then an increase in MHED visits with some fluctuations around the evening hours. Figure B.3 shows the choropleth maps of the number of MHED visits per 1000 sRHA population over 9 fiscal years. Here we can see that sRHAs in RHA 7, the center area of the province, is consistently in the higher range of MHED visits per 1000 sRHA population. Interestingly, the sRHAs in RHA 9, the northern area of the province, is in the lower range of MHED visits in the early 2000s but has an increase in visit rates and is now in the higher range bracket for 2010/2011. These choropleth maps show that the rate of MHED visits is not only different across sRHAs, but over time as well.

The intervals in Figure B.3 were chosen based on Fisher-Jenks natural breaks algorithm (Jenks, 1977) that is based on minimizing the within-class variance. This method was chosen over three other methods: the quantile breaks, equal breaks, and k-means clustering. The equal breaks method divides the data into equally spaced intervals. The quantile breaks method divides the data into intervals such that each interval has the same number of observations. The k-means clustering method uses the algorithm of the same name to cluster the observations based on their similarities to find intervals. To find the best method, or generalizing technique as referred to by Jenks, Jenks developed two techniques, the tabular accuracy index (TAI) and goodness of variance fit measure (GVF) (Jenks and

Caspall, 1971; Jenks, 1977). The TAI is given by

$$TAI = 1 - \frac{\sum_{j=1}^{k} \sum_{i=1}^{N_j} |z_{ij} - \bar{z}_j|}{\sum_{i=1}^{N} |z_i - \bar{z}|} \tag{2.2}$$

and the GVF is given by

$$GVF = 1 - \frac{\sum_{j=1}^{k} \sum_{i=1}^{N_j} (z_{ij} - \bar{z}_j)^2}{\sum_{i=1}^{N} (z_i - \bar{z})^2}. \tag{2.3}$$

Here z_{ij} is the observed value, $i = 1, \ldots, N_j$, $j = 1, \ldots, k$, k is the number of classes, $\bar{z}_j$ is the class mean for class j and N_j is the number of observations in class j. The GVF and TAI range from 0 to 1, where a value of 0 indicates the worst possible fit and a value of 1 indicates the best possible fit (Jenks and Caspall, 1971; Jenks, 1977). For this data, z_{ij} is the number of MHED visits per 1000 sRHA population and the number of classes chosen is 5. The TAI and GVF are shown for the four different generalizing techniques in Table A.4. The Fisher-Jenks natural breaks algorithm has both the highest GVF and TAI thus the intervals in Figure B.3 were chosen based on this algorithm.

2.3 Spatial Autocorrelation

Spatial autocorrelation measures the association between the same variable in "near-by" areas. Cliff and Ord (1973) define spatial autocorrelation as the phenomenon where the presense of some quantity in a region makes its presence in neighbouring regions more or less likely. There are different techniques to measure "nearness" between regions, such as using adjacency or distance between regions. Global spatial autocorrelation measures the overall association of the data. One of the most common global spatial autocorrelation indexes is Moran's I (Moran, 1950), which measures the similarity of rates of contiguous areas. Moran's I is given by,

$$I = \frac{N \sum_{ij} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{(\sum_{ij} w_{ij}) \sum_i (x_i - \bar{x})^2}, \tag{2.4}$$

where N is the total number of observations, x_i and x_j is the quantity of interest at location i and j, respectively, and w_{ij} is the weight assigned to areas i and j. The permutation test approach that uses Monte-Carlo simulation tests the null hypothesis H_0: There is no spatial association in the data. Moran's I assumes that all the observations are independently and identically distributed (i.i.d) with a Gaussian distribution. It assumes that the mean and variance are constant across the regions. However, when the risk population is not constant across regions, this assumption is not met and the permutation test based on Moran's I loses power (Assunçao and Reis, 1999).

The Empirical Bayes Index (EBI) was proposed by Assunçao and Reis (1999) to adjust Moran's I for the variation in population size, but it still assumes constant mean. It has been shown to have a higher power than Moran's I and a stable type I error probability that remained within the nominal significance level in simulated situations (Assunçao and Reis, 1999). The EBI is defined as

$$\text{EBI} = \frac{N \sum_{ij} w_{ij} z_i z_j}{\sum_{ij} w_{ij} \sum_i (z_i - \bar{z})^2}, \tag{2.5}$$

where N is the total number of observations, w_{ij} is the weight assigned to areas i and j and $z_i = \frac{p_i - b}{\sqrt{v_i}}$. Let m_i denote the population size of region i, $m = \sum_i m_i$, and x_i denote the observation at region i, $x = \sum_i x_i$. Here $p_i = x_i/m_i$, the rate in region i, $b = x/m$ and $v_i = s^2 - b/(m/N)$. The sample variance, s^2, is weighted by the population sizes, $s^2 = \sum_i m_i(p_i - b)^2/m$. According to Xiong (2015), the violation of the constant variance assumption is not as serious as the violation of the constant mean assumption. It was found that the violation of the constant mean assumption led Moran's I to incorrectly reject the null hypothesis more often compared to the violation of constant variance. This means the EBI may still lead to an incorrectly rejected null hypothesis and the conclusions must be regarded with caution.

For the PMHC dataset, the maximum and minimum of the sRHA population sizes ranged from 2,236-27,449 and the median was 10,236 in the 2002/2003 fiscal year. There is a large amount of variation between the sRHA population size of Albertans aged 0-17 so permutation tests of the EBI and Moran's I is explored to test for spatial autocorrelation on the global scale. The spatial weights matrix is based on spatial contiguity. In other words, cell (i, j) is equal to 1 if areas i and j are adjacent and $i \neq j$, and 0 otherwise. The spatial weights matrix is row standardized to sum to 1. The results in Table A.5 show that the null hypothesis, no spatial association, is rejected for the fiscal years 2005/2006, 2007/2008, and 2008/2009 at $\alpha = 0.05$. In other words, there is statistical evidence that some positive spatial autocorrelation exists for these fiscal years. Though their statistics and p-values are different, the conclusions are the same for the global Moran's I and EBI.

To study the contribution of each observation to the global Moran's I, Anselin (1995) developed local indicators of spatial association, abbreviated to LISA. The formula for the local Moran's I is

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1)}, \tag{2.6}$$

where the sum of the local Moran's I's for all observations is proportional to the global Moran's I. These indicators serve two main purposes. The first is to identify local spatial clusters, or "hot spots", and the second is to assess the influence of individual locations on the magnitude of the global index and to identify potential outliers (Anselin, 1995). The

null hypothesis is that there is no association between the value observed at location i and the values observed at nearby sites, determined through w_{ij}. LISAs using the number of MHED visits in sRHA i over the population size of sRHA i are calculated for each of the nine fiscal years. The results found significant LISAs for sRHAs in every RHA aside from R2. For an sRHA in R1, it was found to be significant for the fiscal years 2007/08 to 2008/09 and sRHAs in R3 were significant for fiscal years 2002/03 to 2005/06 and 2009/10. Some sRHAs in R4 had significant LISAs for four of the fiscal years, R6 had some sRHAs with significant LISAs for three fiscal years and R7 had sRHAs with significant LISAs the most frequently. The sRHAs in R7 had significant LISAs for every fiscal year aside from 2002/03 and 2003/04. In RHA R8, it had significant LISAs for six of the fiscal years. There was one influential sRHA in R9 that had a significant LISA in 2008/09.

The final spatial autocorrelation analysis examined is the spatial autocorrelation of the sRHAs within their respective RHAs. To do this, a global Moran's I is calculated for each RHA for the nine fiscal years. The results found that R1 rejected the null hypothesis of no spatial autocorrelation for all but fiscal year 2007/08. In addition, R4 and R7 rejected the null hypothesis for fiscal year 2010/11 and R6 rejected the null hypothesis in 2007/08.

The results indicate that there may be some presence of spatial autocorrelation at the global level and some indication of clustering at the local level. Globally there is a presence of spatial autocorrelation for three of the nine fiscal years and locally we found the sRHAs in R7 to have potential clusters the most frequently. The results also indicate that there is a presence of global spatial autocorrelation for the sRHAs within RHAs R1, R4, R6 and R7. When the model is being constructed, a component for the spatial correlation should be considered.

2.4 Temporal Autocorrelation

Temporal autocorrelation is the correlation of a variable with itself over time. In Figure B.4, there appears to be a cyclical pattern to the autocorrelation for the majority of the RHA ACFs, with strong positive autocorrelation at intervals of 12 months and multiples thereof, and matching negative correlation at 6 months, 18 months, etc. For RHA 9, there is no pattern in the autocorrelations and no strong presence of temporal autocorrelation. The cyclical pattern is most prominent for RHAs 3 and 6, the Calgary and Edmonton areas, respectively. These plots strengthen the argument to include a cyclic seasonal effect in the model. In Figure B.5, the majority of the RHAs aside from R7, R8 and R9 have some significant lag effects up to lag 20 months and no significant lags afterwards. The figure shows that the autocorrelation is not as severe after approximately lag 20.

Chapter 3

Generalized Linear Mixed Models (GLMM)

Chapter 2 found that when the spatio-temporal effects are analyzed separately there is a presence of spatial autocorrelation, temporal autocorrelation and a cyclic seasonal effect. Chapter 3 introduces the proposed models to explore these spatio-temporal patterns. Section 3.1 begins by introducing some notation. Then Section 3.2 presents the proposed models divided into two modelling stages: temporal and spatio-temporal. The models explore three different specifications for the seasonal and time trend effect and random effects to model the spatio-temporal correlation structure of the data.

3.1 Notation

Let $Y(t; r; \boldsymbol{p})$ denote the primary response where t is the time index, r is region index and $\boldsymbol{p}$ is the covariate vector index. $Y(t; r; \boldsymbol{p})$ can be the number of MHED visits, the number of subjects who had an MHED visit, a ratio of the number of MHED visits over population sizes or a ratio of the number of subjects with an MHED visit over the population sizes. This project focuses on the response $Y(t; r; \boldsymbol{p})$ seen as the number of MHED visits at time t and region r with covariate information $\boldsymbol{p}$. An MHED visit constitutes those made by Albertans aged 0-17 at the time of their MHED visit.

For the region index, r, of this project, it is treated as a discrete index such that $r = 1, \ldots, 70$ for the 70 sRHAs. The sRHAs correspond to the smallest regional index of this data. The region index can also be treated as centroids with latitude and longitude values. The calendar time of April 1, 2002 to March 31, 2011, t, can be in yearly, monthly, weekly or daily time. For this project, we instead look at time as 28 day blocks starting at April 1, 2002, giving us 13 blocks each year. The index t then goes from $t = 1$ to $T = 9 \times 13 = 117$. The advantage of using 13 blocks over monthly time is that each time unit has an equal amount of days. This ensures the variations in counts between blocks will not be due to differences in amount of days. The covariate vector index, $\boldsymbol{p} = (j, k, l)$, is a vector of indices for age,

gender, and pSES, respectively. Here, $j = 1, 2$ for age groups $0-12$ and $13-18$, $k = 1, 2$ for genders female and male and $l = 1, 2, 3$ for pSES categories regular plan participant (RPP), treaty status, and government sponsored and welfare combined. We refer to the combined grouping of government sponsored and welfare as income supported.

Let $S(t; r)$ denote the population count at time t and sRHA r of Albertans aged 0-18. The true population count at time t and region r is unknown for most time points as it is only recorded for the sRHA level at the fiscal year end. For the purposes of this project, we assume the population remains constant throughout the fiscal year. The goal is to explore different methods to specify and estimate the conditional expectation of the response $Y(t; r; \boldsymbol{p})$ given the population:

$$\mathrm{E}\{Y(t; r; \boldsymbol{p})|S(t; r)\} = \mu(t; r; \boldsymbol{p}|S(t; r)). \tag{3.1}$$

We denote the conditional covariance as

$$\mathrm{Cov}\{Y(t_1; r_1; \boldsymbol{p}_1), Y(t_2; r_2; \boldsymbol{p}_2)|S(t_1; r_1), S(t_2; r_2)\} = c(t_1, t_2; r_1, r_2; \boldsymbol{p}_1, \boldsymbol{p}_2|S_1, S_2). \tag{3.2}$$

3.2 Modelling

To model $\mu(t; r; \boldsymbol{p}|S(t; r))$, the problem is separated into two modelling stages: temporal and spatio-temporal. This is to explore the spatio-temporal patterns separately and build the final model appropriately. The temporal modelling considers three different settings, a log-linear, non-linear, and semi-parametric setting, for the seasonal and time trend effect. The spatio-temporal modelling considers random effects to describe the nested structure of the regions followed by CAR and AR(1) random effects.

We decompose the model into three components: the risk factors/exposure, the seasonal effect and time trend, and the random effects. The exposure refers to the population size $S(t; r)$. Since the response variable is in the form of counts, the Poisson distribution is assumed. We denote the Poisson regression model as

$$\log(\mu(t; r; \boldsymbol{p}|\boldsymbol{\varepsilon})) = \boldsymbol{X}\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t; r) + \boldsymbol{\beta}(t; r) + \boldsymbol{\varepsilon}(t; r), \tag{3.3}$$

where $\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t; r)$ denotes the risk factors/exposure fixed effects, $\boldsymbol{\beta}(t; r)$ denotes the seasonal and time trend fixed effects and $\boldsymbol{\varepsilon}(t; r)$ denotes the random effects component. The specification of $(t; r)$ denotes whether the components vary by t and/or r. We denote $\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t; r) = (\gamma(t; r), \theta_0(t; r), \alpha_j^{\text{age}}(t; r), \alpha_k^{\text{gender}}(t; r), \alpha_l^{\text{pSES}}(t; r))$ and $\boldsymbol{X}$ as the design matrix of $\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t; r)$. The first column of the design matrix is $\log S(t; r)$, the second column is a column of ones for the intercept and the remaining four columns correspond to the categorical risk factors. Let $\alpha_1^{\text{age}}(t; r) = 0$, $\alpha_1^{\text{gender}}(t; r) = 0$ and $\alpha_1^{\text{pSES}}(t; r) = 0$. In other

words, individuals aged 0-12 years old, females and pSES category RPP are the reference groups of the risk factors.

3.2.1 Modelling temporal patterns

We begin by considering three methods to describe the seasonal and time trend effect, $\boldsymbol{\beta}(t;r)$. This effect is considered under fixed regions, r^*, such that $\boldsymbol{\beta}(t;r^*)$. In other words, the seasonal effects are estimated for each r^* region. To see whether this specification is appropriate, the seasonal effects varying across regions, $\boldsymbol{\beta}(t;r^*)$, are compared to seasonal effects that are constant across all regions. When the seasonal effects are considered under fixed regions, the risk factors/exposure is $\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t;r^*)$. For this section we assume there are no random effects, $\boldsymbol{\varepsilon}(t;r^*) = 0$.

Model T1 assumes a linear time trend and assigns a coefficient for each of the 12=13-1 blocks of size 28 days such that

$$\boldsymbol{\beta}(t;r^*) = \phi(t;r^*)x_{tr^*\boldsymbol{p}} + \beta_{i(t)}^{\text{season}}(t;r^*). \tag{3.4}$$

Let x_{tr^*p} denote the time at t and region r^* for risk groups $\boldsymbol{p}$. Let $i = 1,\ldots,13$ for the 13 blocks of size 28 days in a year and let block 1 be the reference group, $\beta_{1(t)} = 0$.

Model T2 assumes a linear time trend and a seasonal pattern that is described by sine and cosine functions such that

$$\boldsymbol{\beta}(t;r^*) = \phi(t;r^*)x_{tr^*\boldsymbol{p}} + A_0(t;r^*)\cos\left(\frac{2\pi t}{13}\right) + B_0(t;r^*)\sin\left(\frac{2\pi t}{13}\right). \tag{3.5}$$

Let x_{tr^*p} denote the time at t and region r^* for risk groups $\boldsymbol{p}$.

Model T3 assumes a non-linear time trend and a cyclic seasonal pattern using non-parametric smoothing functions such that

$$\boldsymbol{\beta}(t;r^*) = f_{\text{trend,t}}(t;r^*) + f_{\text{seasonal,i(t)}}(t;r^*), \tag{3.6}$$

where i and t are covariates that describe the within-year and between year times, respectively. In other words, $i = 1,\ldots,13$ for the 13 blocks of size 28 days in a fiscal year and $t = 1,\ldots,117$ for the 13×9 total time points. The smoothing functions $f_{\text{seasonal,i(t)}}$ and $f_{\text{trend,t}}$ use a cyclic smoothing spline and thin plate smoothing spline, respectively. The cyclic cubic spline is defined as $f_{\text{seasonal,i(t)}} = \sum_{i=1}^{13-1} \tilde{b}_{i(t)}(i)\beta_{i(t)}$, where $\tilde{b}_{i(t)}$ are the basis functions. A characteristic of the cyclic cubic spline is that $f_{\text{seasonal,1(t)}}$ must match $f_{\text{seasonal,13(t)}}$, up to a second derivative (Wood, 2006). The cyclic cubic smoothing spline can be seen as an extension of Model T1 with 13-1 seasonal factors where the spline is applied to smooth over the seasonal effect. The thin plate spline provides intuition on whether a linear time trend is adequate or whether a polynomial time trend is required.

3.2.2 Modelling spatial-temporal patterns

This section begins by attempting to use the risk factors and seasonal effect chosen through the temporal modelling stage to describe the spatial patterns. Recall that the notation t^* and r^* refers to estimated for fixed time t^* and estimated for fixed region r^*. Here β^{seasonal} refers to the seasonal component of $\boldsymbol{\beta}$.

Model ST1 assumes the fixed effects are constant across time and space, the time trend component of $\boldsymbol{\beta}$ is zero and there are no random effects. This gives us the following model specification,

$$\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}; \qquad \boldsymbol{\beta} = \beta^{\text{seasonal}}; \qquad \boldsymbol{\varepsilon} = 0. \tag{3.7}$$

Model ST2 assumes the fixed effects vary across time but are constant across space. We assume the time trend component of $\boldsymbol{\beta}(t^*; r)$ is zero and there are no random effects. In other words, it extends model ST1 to let the coefficients vary across time by estimating the fixed effects for each time unit. To investigate how the seasonal effect varies over time, the fixed effects vary across the nine fiscal years rather than all $t = 1, \ldots, 117$ time points. This gives us the following model specification,

$$\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t^*_{\text{year}}; r); \qquad \boldsymbol{\beta}(t^*_{\text{year}}; r) = \beta^{\text{seasonal}}(t^*_{\text{year}}; r); \qquad \boldsymbol{\varepsilon}(t^*_{\text{year}}; r) = 0, \tag{3.8}$$

where $t^*_{\text{year}} = 1, \ldots, 9$ for the nine fiscal years.

Model ST3 is a generalized linear mixed model, GLMM, with independent random intercepts for the RHA and sRHAs. This is to describe the hierarchical structure of the data, where the sRHAs are nested within their respective RHAs. This structure assumes the sRHAs in the same RHA to be correlated and the sRHAs not within the same RHA are assumed to be independent. The random effects specification is denoted as

$$\boldsymbol{\varepsilon}(t; r) = u_R + v_r, \tag{3.9}$$

where $R = 1, \ldots, 9$ for the nine RHAs, $r = 1 \ldots, 70$ for the 70 sRHAs, $u_R \overset{iid}{\sim} N(0, \sigma_u^2)$ and $v_r \overset{iid}{\sim} N(0, \sigma_v^2)$. The random effects u_R and v_r are assumed to be independent. The temporal fixed effects component of the model will be chosen through the temporal analysis and the risk factors/exposure component will vary by time and/or regions depending on the analysis results.

Model ST4 assumes a temporal and spatial random effect. The goal of this final model is to incorporate the findings from the previous models and to describe the correlation structure. The seasonal and time trend, $\boldsymbol{\beta}(t; r)$, and risk factors/exposure, $\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exposure}}(t; r)$, fixed effects specification of varying by time and/or regions will be based on the analysis

results in Chapter 4. The random effect specification is

$$\boldsymbol{\varepsilon}(t;r) = \psi_t + \eta_r, \tag{3.10}$$

where $\boldsymbol{\psi}$ follows an autoregressive model of order 1, AR(1), and η_r follows a conditionally autoregressive, CAR, model. Let $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_T) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\psi)$ where $\Sigma_\psi = \sigma^2_\psi/(1 - \rho^2)\{\rho^{|t_1-t_2|}\}^T_{t_1,t_2=1}$. The parameter σ^2_ψ is the temporal variance and ρ is the lag-1 temporal autocorrelation. Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{70}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$ where

$$\boldsymbol{\Sigma}_\eta = \sigma^2_\eta(\mathbf{I}_{70} - \lambda_\eta \mathbf{W})^{-1}. \tag{3.11}$$

Here σ^2_η is the spatial dispersion, the scalar parameter λ_η describes the strength of spatial dependence among the observations (LeSage and Pace, 2009), $0 \leq \lambda_\eta < 1$, and $\mathbf{W}$ is the spatial weights matrix. The spatial weights matrix is defined as the adjacency matrix where if two different regions are adjacent, the corresponding matrix element will be 1 and 0 otherwise. This model allows for correlation amongst sRHAs who share a common border, whereas the sRHA and RHA random effects in model ST3 allowed the sRHAs within the same RHA to be correlated. The CAR model focuses on spatial correlation based on the geographic location and the nested random effects in model ST3 focuses on correlation based on the geographic hierarchical structure created by the government. Model ST4 uses an AR(1) and CAR model similar to the one described in Torabi and Rosychuk (2010) and the CAR model described in LeSage and Pace (2009).

Chapter 4

Regression Analysis Under GLMM

In this chapter, we begin by outlining the analysis procedure in Section 4.1. Sections 4.2 and 4.3 summarize the analysis for the temporal and spatio-temporal modelling stages detailed in Chapter 3. The final proposed model is chosen based on the results. For clarification, Table A.6 shows the level of aggregation at the sRHA level used for the analysis. There are 70 sRHAs, 117 (9 fiscal years $\times$ 13 blocks of size 28 days) time points, 2 levels of gender (Male and Female), 3 levels of pSES (Regular Plan Participant (RPP), Treaty Status, and Income Supported), and 2 age groups (less than 13 years old and greater than 12 years old). The distribution of the MHED counts following this data aggregation and the distribution of the population counts at the sRHA level are shown in Figure B.6.

4.1 Overall Analysis Procedure

For each of the proposed models, the data are based on the number of MHED visits in the subgroup $(t; r; \boldsymbol{p})$. The analysis begins with the temporal modelling and assumes the coefficients are constant across all regions. To implement the three proposed models, models T1 and T2 uses the standard `glm` function in `R`. The final model, T3, is fit using the `gam` function in the `mgcv` `R` package where the smoothing functions are specified to be the cyclic cubic spline with basis dimension $k = 13 - 1$ and the thin plate spline for the seasonal and trend functions, respectively. The default smoothing parameter estimation method in the `R` package is used, generalized cross validation (GCV). Next, the coefficients are assumed to vary by RHA and sRHA for models T1, T2 and T3. The estimation by region is done by subsetting the data for each RHA and sRHA and running a separate model on each subset of the data. Another option is to include an interaction term of the regions, treated as a categorical variable, and the fixed effects model components. The intercept must be removed for this option so it can be replaced by the means of the regions. To determine whether the coefficients should vary by region, the confidence intervals (CIs) of the estimates are compared to their constant across regions CI counterpart. If they fall within the constant estimate CI, the constant estimate is used. Otherwise, some form of grouping of the regions

has to be explored. In addition, if the γ coefficient for population size includes one in its 95% CI, it is fixed at one and treated as an offset term.

The analysis for the spatio-temporal modelling starts by implementing models ST1 and ST2, the constant and time-varying Poisson GLM, respectively. For model ST2, the model is fit with the data divided by fiscal year. If the estimated coefficients fall inside their respective constant coefficient estimate CI from model ST1, the variable is assumed to be constant across time. The mixed effects model, ST3, with a nested random effect structure is fit with `glmer` in the `R` package `lme4`. Like in the temporal modelling stage, if the coefficient for population size, γ, includes one in its CI, it is fixed at one and treated as an offset. Lastly, the final model ST4 will use the information gained from the previous models. If one method outperforms the others in its respective stage, that method is the focus for this final model. To implement the CAR and AR(1) random effects, the integrated nested Laplace approximation using the `R` package `INLA` is used (Rue et al., 2009).

4.2 Temporal Analysis

Recall that three different models were proposed to capture the seasonality and time trend: Model T1 that estimates a coefficient for 13-1 blocks of size 28 days and a linear time trend, Model T2 that uses trigonometric functions and a linear time trend and Model T3 that uses a cyclic cubic spline for the seasonality and a thin plate spline for the time trend. The demographic, seasonal and time trend coefficients are proposed to be estimated under three different settings for each of the three models. The first setting fixes the coefficients to be constant for every region. The second setting allows the coefficients to vary by RHA and the third allows the coefficients to vary by sRHA. To achieve this, subsets of the data by RHA and sRHA for settings 2 and 3, respectively, are fit to the models. These settings help see whether there is a difference between regions and if there is a loss of information by ignoring these differences.

4.2.1 Risk factors/exposures effects

To begin, we compare the coefficient estimates for the risk factors and exposures across the three different settings, constant across regions and varying by RHA and sRHA. In Figure B.7, we see the coefficients for the intercept, population, and gender estimated by sRHA, RHA and constant across all regions. The x-axis is the sRHA index and the y-axis is the coefficient's estimated value. The intercept and population coefficient have similar results for which sRHAs estimates fall inside the RHA or constant estimate's confidence interval (CI). Aside from R303, R313, R505, R611, R620, R901 and R902, all of the sRHAs fall inside the constant and RHA coefficient estimate CIs for the intercept and population variable. The gender coefficient estimates by sRHA fall inside the constant estimate CI aside from sRHA R202 and sRHAs in RHAs R3 and R9. The estimates by sRHA all fall inside the

RHA estimates' CI. This motivates estimating the gender coefficient by RHA or groups of RHAs.

In Figure B.8, we see the teenage age group coefficient estimates varying by sRHA all fall inside the RHA estimate CI aside from R902. A large portion of the teenage coefficient estimates by sRHA in RHAs R3, R6 and R9 do not fall inside the constant estimate CI. Like the gender coefficient, the age group coefficient may be more informative if estimated by some grouping of the RHAs. Figure B.8 shows that none of the treaty status coefficient estimates varying by sRHA, compared to RPP, fall within the constant estimate's CI. The sRHA coefficient estimates in RHA R5 fall inside the RHA estimate CI. For the income supported coefficient estimates varying by sRHA, compared to RPP, the estimates for the Calgary area and Edmonton area RHAs R3 and R6, respectively, do not fall inside either the constant or RHA estimate CI. The majority of the estimates by sRHA in RHA R9 fall inside the constant and RHA estimates' CI. The remaining estimates by sRHA fall inside their respective RHA estimate CI.

The results indicate that the sRHAs belonging to the same RHA do not always follow the same pattern, particularly when looking at pSES. Treating them as such may cause us to lose information. The sRHAs that tend to follow a different pattern from the overall average are from RHAs R3, R6 and R9. This intuitively makes sense since R9, the northern region of Alberta, has sparse population and may be influenced by R904, Fort McMurray, which accounts for over 57% of the RHAs population. In addition, R3 and R6, the Calgary and Edmonton area, respectively, have the largest amount of sRHAs, 19 and 18, respectively. There is a large amount of variability between the population sizes of the sRHAs in R3 and R6. For R3, in 2003 the standard deviation of the sRHAs population sizes was 5,612.3, and in 2011 that increased to 8,223.5. For R6, in 2003 the standard deviation of the sRHAs population sizes was 7,567.2, and in 2011 that increased to 9,022.2. Given these results, it is recommended the coefficients be estimated under four groups: RHA 3, 6, 9 and the remaining RHAs grouped into one category "Other".

4.2.2 Estimates of average counts

All three temporal models estimated the average MHED counts for individuals aged less than 13 to be less than one regardless of the time, sRHA, pSES and gender since over 92% of this age group has no MHED visits. In addition, of those in that age group who had an MHED visit, 85% had a count of one visit. For the estimated means of those aged greater than 12 and in pSES group treaty status, the constant, varying by RHA and sRHA settings only differed for a few sRHAs in RHA R1, R4, R6 and R7 for a total of nine sRHAs. This means that for teenagers with treaty status, only nine sRHAs have different behaviour compared to their RHAs and Alberta as a whole. For those aged greater than 12 and in pSES group income supported, the estimated means are similar aside for some sRHAs in R6 (Edmonton area). For individuals aged greater than 12 and in pSES category

RPP, the waves have a larger magnitude when the coefficients vary by sRHA. Overall, the estimated means varying by RHA and sRHA have a similar time trend and estimated means. Based on these results, the model may benefit from allowing the pSES variable, time trend and seasonal effect to vary by RHA. Figure B.9 shows the estimated means by 1000 sRHA population for the three different pSES categories for females and individuals aged greater than 13. Notice how the coefficients estimated by RHA and sRHA are similar and capture which sRHAs display a different behaviour from the other sRHAs better than the coefficients held constant.

When the estimates are compared across models T1, T2 and T3, the seasonal factors model, T1, closely resembles the smoothing splines model, T3. Model T3 estimates have a non-linear time trend and the seasonality is smoother compared to the seasonal factors model. The magnitude of the seasonal waves in the estimated means is smaller for the trigonometric model, T2, compared to models T1 and T3. This is seen in Figure B.10. The difference between the models is less severe when the coefficients are estimated by sRHA. Note the spikes some of the curves have at the beginning of the fiscal years in Figure B.9, particularly for the coefficients estimated by sRHA. This is most prominent in the third column and second row plot, the trigonometric model for treaty status female teenagers in sRHA R703. This is due to the estimate for the log of the population size having a larger range when it was estimated by sRHA, -16 to 73, compared to the coefficients estimated by RHA, -0.47 to 1.57. The jagged shape of the curve is also a side effect of the assumption that the population size remains constant throughout the fiscal year. The change in population size at the start of the fiscal year compared to the previous year causes a jump in the estimated means.

4.2.3 Seasonality and time trend

The results have indicated that the trigonometric model, T2, is a good model to capture the seasonal variation in the estimated mean number of MHED visits. This conclusion is supported by models T1 and T3, the seasonal factors and smoothing splines models. In model T1 a seasonal variation similar to the cosine + sine curve was observed. This is shown in Figure B.11 where we see the seasonal factor follows a cosine + sine curve for all of the regions. Figure B.12 shows the estimated time trend + the seasonal effect from models T1 and T2 calculated for $t = 1, \ldots, 117$, each RHA and the Alberta average. The magnitude of the waves in the trigonometric model is smaller than the magnitude in the seasonal factors model, but the seasonal patterns are similar. In Figure B.13, the spline describing the block effect shows a curve similar to the one found in the trigonometric model. The figures also show that the assumption of a linear time trend for time t may not be appropriate for all RHAs and polynomial terms for time should be considered.

4.2.4 Model checking

The format of the data is a 5-way contingency table (sRHA by time by gender by age group by pSES) to ensure a balanced design. Due to the large number of combinations, 98,280, there is a large amount of zeros in the data, approximately 77%. This motivates the exploration of different distributions aside from the Poisson distribution that are more suited for zero-inflated data such as the negative binomial or the quasi-Poisson.

In a Poisson GLM, the dispersion parameter is fixed at 1. The quasi-Poisson distribution leaves the dispersion parameter unrestricted and is estimated using the data. The coefficient estimates of the quasi-Poisson remain the same as the Poisson GLM but the standard errors in the quasi-Poisson are adjusted for over-dispersion. For the three proposed models, the seasonal factors, trigonometric and generalized additive models, the dispersion parameter is less than 2 under all three settings where the coefficients varied by sRHA, RHA and were constant across all regions. The results for the dispersion parameter are displayed in Table A.7. Given the estimated dispersion parameter is less than 2 and approximately greater than 1 for all models, aside from two sRHAs, over-dispersion does not appear to be a concern for our Poisson regression models.

Another method of handling over-dispersed, or zero-inflated data, is to use the negative binomial distribution. In Table A.8, the estimates and standard errors are very similar for the negative binomial distribution and the Poisson distribution for models T1, T2 and T3 when the variables are fixed over time and regions. Their distribution of residuals vs fitted values were approximately the same under the negative binomial and Poisson distribution. A measure of the model's fit is the deviance explained. It is a measure of how well the explanatory variables in the model explain the variation in the response and is calculated as 1-residual deviance/null deviance. The deviance explained of the three proposed models is slightly higher under the negative binomial distribution compared to the Poisson distribution. For the Poisson and negative binomial distribution, the deviance explained is approximately 40% under the constant coefficients setting. Under the varying by RHA coefficient setting the deviance explained is on average 43% ($\pm$ 8%) and under the sRHA coefficient setting the deviance explained is on average 44% ($\pm$10%). These results indicate the negative binomial regression models do not provide a significantly better fit over the Poisson regression models.

4.2.5 Analysis of residuals

Figure B.14 shows the residual versus fitted value plots for (a) the constant coefficients across regions, (b) the coefficients varying by RHA, and (c) the coefficients varying by sRHA settings for Model T2. From the plots, the coefficients varying by sRHA appears to provide a slightly better fit compared to the coefficients varying by RHA due to there being

fewer outliers for the fitted values close to zero. The constant coefficients across regions setting has a smaller range for the fitted values and is more overwhelmed by the zeroes.

The residuals can be used to check if there is any remaining temporal autocorrelation not described by the model. If we have described the temporal pattern sufficiently, there should be no presence of temporal autocorrelation. The results described in the following paragraph are the same for models T1, T2 and T3. They are also true for the coefficients estimated under the three settings of remaining constant across all regions, varying across RHAs and varying across sRHAs.

If we treat each combination of sRHA, gender, pSES and age group as a time series with 13 blocks $\times$ 9 fiscal years = 117 time points, we have $70 \times 2 \times 3 \times 2 = 840$ time series. There was one sRHA, R614, for females and the age group 0-12 years old that had a presence of autocorrelation for pSES category RPP. Aside from that one time series in pSES category RPP, there was no presence of autocorrelation for all combinations of gender, age group and pSES category RPP. For the pSES category treaty status, there is still a presence of autocorrelation for the majority of sRHAs. The presence of autocorrelation is less severe for treaty status females aged 13-18 years old. For the pSES category income supported, individuals aged less than 13 still have a presence of autocorrelation for several sRHAs. For individuals aged greater than 12, only females in R620 have a presence of temporal autocorrelation. These results indicate that an interaction between pSES and time should be included to try and capture that remaining temporal autocorrelation.

4.2.6 Proposed model given temporal analysis

To capture the seasonality in the MHED counts, the trigonometric model T2 with a cosine and sine function is recommended over the cyclic cubic spline and the 13-1 factors for the blocks of size 28 days. This is due to the trigonometric model having fewer parameters compared to the seasonal factors model and the trigonometric model being more interpretable compared to the smoothing splines. In addition, we continue to assume the data follows a Poisson distribution due to the estimated over-dispersion being relatively small. For the risk factors, we allow them to vary based on 4 groupings of the RHAs: R3, R6, R9 and Other. Due to the pSES categories still experiencing some temporal autocorrelation, an interaction term of pSES and time t is recommended. The smoothing splines revealed a linear trend may not be appropriate. Given these results, a quadratic term for t is recommended.

4.3 Spatio-Temporal Analysis

For the spatio-temporal modelling in Section 3.2.2, recall that four models were proposed: model ST1 is a GLM with coefficients constant over time, model ST2 is a GLM with coefficients varying over yearly time, model ST3 is a mixed effects model with a nested random effect for the sRHA nested within the RHAs and model ST4 is a mixed effects model with

AR(1) and CAR model random effects. For models ST1 and ST2, the trigonometric seasonal effect is used and the coefficients are estimated by fiscal year. The results are summarized in Table A.9. Aside from fiscal year 2003/04, all of the estimates for the seasonal effect, A_0 and B_0, fall inside the constant estimate's CI. For the coefficients for male and teenagers, all of the estimates fall within the constant estimate's CI. This means the coefficients for the trigonometric seasonal effect, gender and age do not experience significant variation over yearly time. For the intercept coefficient, fiscal years 2003/04, 2006/07 and 2008/09 do not fall within the constant estimate's CI. For the log of the population coefficient, fiscal years 2003/04 and 2008/09 don't fall within the constant estimate's CI. Since the deviation from the constant estimate's CI is small, less than 0.08, for all but the intercept estimate in fiscal year 2003/04, the constant estimate is recommended for the intercept and log of the population in favour of a sparser model. For the pSES categories treaty status and income supported, fiscal years 2002/03, 2007/08, 2009/10 and 2010/11 do not fall within the constant estimate's CI. This provides further motivation for including an interaction term of time and pSES.

The third model for the spatio-temporal modelling stage, ST3, is a mixed effects model that incorporates the nested structure of the regions. To formulate this model, information from all previous models is used and the structure from Equation 3.3 holds. The fixed effects for the risk factors/exposure are

$$\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exp}}(t; R^{*,\text{group}}) = (\gamma, \theta_0, \alpha_j^{\text{age}}(t; R^{*,\text{group}}), \alpha_k^{\text{gender}}(t; R^{*,\text{group}}), \alpha_l^{\text{pSES}}(t; R^{*,\text{group}})), \quad (4.1)$$

where the intercept and population coefficient, θ_0 and γ are constant across time and space, the risk factors age, gender and pSES vary by regions $R^{*,\text{group}} = R3, R6, R9$ and the other RHAs. The temporal fixed effects coefficients are estimated to be constant across time and regions such that

$$\boldsymbol{\beta} = A_0 \cos\left(\frac{2\pi t}{13}\right) + B_0 \sin\left(\frac{2\pi t}{13}\right) + \phi_0 x_{tr\boldsymbol{p}} + \phi_{0l}^{\text{pSES}} x_{tr\boldsymbol{p}} + \phi_2 x_{tr\boldsymbol{p}}^2, \quad (4.2)$$

where a quadratic term is included for time through ϕ_2. The term ϕ_{0l}^{pSES} is the interaction of time and pSES. The random effects are $\boldsymbol{\varepsilon}(t; r) = u_R + v_r$ where $R = 1, \ldots, 9$ for the RHAs, $r = 1, \ldots, 70$ for the sRHAs, $u_R \sim N(0, \sigma_u^2)$, $v_r \sim N(0, \sigma_v^2)$ and u_R and v_r are independent.

Recall that in Figure B.12 the seasonality and time trend varied across the RHAs. To explore this seasonal and time trend effect further, consider the above model under three settings for the seasonal and time trend effect.

1. Assume the coefficients are constant: $\boldsymbol{\beta}$ from Equation 4.2.

2. Assume the coefficients vary by RHA: $A_0(t; R^*)$, $B_0(t; R^*)$ and $\phi_0(t; R^*)$.

3. Assume a mixed effects model where the seasonal and time trend coefficients are assumed to vary randomly by RHA: $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}(t;r) = u_R + v_r + A_{1R}\cos\left(\frac{2\pi t}{13}\right) + B_{1R}\sin\left(\frac{2\pi t}{13}\right) + \phi_{1R}x_{trp}$ such that $A_{1R} \sim N(0,\sigma_A^2)$, $B_{1R} \sim N(0,\sigma_B^2)$ and $\phi_{1R} \sim N(0,\sigma_\phi^2)$. Assume A_{1R}, B_{1R} and ϕ_{1R} are correlated random effects.

The AICs and BICs for these models are displayed in Table A.10. Based on the AICs and BICs, the model with the random seasonal and time trend effect offers the best fit compared to the models that left these effects fixed. Figure B.15 shows the residual versus fitted value plots for the generalized linear mixed models under the three settings. Plot (a) assumes the seasonal and trend coefficients are constant, plot (b) assumes the coefficients vary by RHA and plot (c) assumes the coefficients are random based on the RHA groups. There is no discernible difference between these plots. Given these results, the model should include random effects for the seasonal and time trend slopes. The variance estimate for σ_u^2, the variance of the RHAs, is small, < 0.01. This means the correlation between sRHAs in the same RHA may be insignificant. This will be explored further in the analysis of model ST4.

Current results have indicated that the number of MHED visits varies over time and space. Currently the suggested model is a random slope for the trigonometric seasonal effect and time trend varying by RHA, a nested random effect of the sRHA nested within the RHA and risk factors/exposures that vary according to four groups, RHA 3, RHA 6, RHA 9 and the remaining RHAs. The risk factors/exposure and temporal fixed effects remain the same as in Equation 4.1 and 4.2, respectively. The random effects structure is now $\boldsymbol{\varepsilon}(t;r) = u_R + v_r + A_{1R} + B_{1R} + \phi_{1R}$. Here $R = 1,\ldots,9$ for the nine RHAs, $r = 1,\ldots,70$ for the 70 sRHAs, $u_R \sim N(0,\sigma_u^2)$, $v_r \sim N(0,\sigma_v^2)$, $A_{1R} \sim N(0,\sigma_A^2)$, $B_{1R} \sim N(0,\sigma_B^2)$ and $\phi_{1R} \sim N(0,\sigma_\phi^2)$. A final model to consider, model ST4, is one that includes a spatial and temporal correlation structure, such as the CAR and AR(1) model, respectively. To implement this model we use the `R` package `INLA` that performs approximate Bayesian inference.

4.3.1 Motivation for Integrated Nested Laplace Approximation (INLA)

The `R` package `INLA` uses an integrated nested Laplace approximation approach to approximate Bayesian GLMs. It uses nested approximations for the posterior marginals of interest and applies a Laplace approximation to the nested approximations (Rue et al., 2009). A Bayesian GLM is commonly denoted as

$$g(\mu_i) = \eta_i = \alpha + \sum_{k=1}^{n} \beta_k z_{ki} + \sum_{j=1}^{n_f} f^j(u_{ji}), \tag{4.3}$$

where g is the link function, μ_i is the mean, f^js are unknown functions of the covariates $\boldsymbol{u}$ and β_ks are the fixed effects of the covariates $\boldsymbol{z}$. Latent Gaussian models are a subset of these Bayesian models which assign a Gaussian prior to α, f^j and β_k. The latent field $\boldsymbol{x}$ is

a vector of the latent Gaussian variables η_i, α, f^j and β_k. The latent field $\boldsymbol{x}$ has conditional independence properties that makes it a Gaussian Markov Random Field (GMRF) (Rue et al., 2009). If we have $\boldsymbol{x} = (x_1, \ldots, x_n)' \sim N$, then $x_i \perp x_j | \boldsymbol{x}_{-ij}$, in other words, x_i and x_j are conditionally independent. This key property leads to a sparse precision matrix $\boldsymbol{Q}$, the inverse of the covariance matrix, leading to fast computation speed (Rue et al., 2009).

This method is chosen due to the ability to specify different models for the latent Gaussian models such as the AR(1) model, CAR model and independent random variables. Another method would be to use the `R` function `glmmPQL` that uses penalized quasi-likelihood (Wolfinger and O'Connell, 1993). However according to Bolker (2015), the estimates of the random-effects variances are biased, particularly for binary data or count data with means less than 5. Our response has a mean of 0.42 so this method is not ideal. In addition, since the quasi-likelihood is computed, the inference is limited. The second option is to use the relatively new `R` function `glmmTMB`. The function has the option of including an AR(1) model for the correlation structure. However, it is not used due to the documentation on how this AR(1) structure is implemented being unclear and due to a CAR model correlation structure not currently being available.

4.3.2 Model exploration

Currently the model includes interaction terms for the demographic covariates with time and space. There could exist an interaction between the demographic variables themselves. Newton and Rosychuk (2011) found that the rate of MHED visits was approximately equal for males and females in the younger age groups. They found that for the older age groups of 15-17, females were nearly twice as likely to visit an ED for mental health reasons compared to males. This motivates the inclusion of an interaction term for age and gender. An interaction term for age and pSES is included since it is found to be significantly different from zero. The interaction term for gender and pSES was not found to be significant so it is excluded from the proposed models. The temporal fixed effects remain the same as in Equation 4.2 and for the risk factors/exposure with the two-way interactions we have

$$\begin{aligned}\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exp}}(t; R^{*,\text{group}}) = \Big(&\gamma, \theta_0, \alpha_j^{\text{age}}(t; R^{*,\text{group}}), \alpha_k^{\text{gender}}(t; R^{*,\text{group}}), \alpha_l^{\text{pSES}}(t; R^{*,\text{group}}), \\ &\alpha_{jk}^{\text{age}\times\text{gender}}, \alpha_{jk}^{\text{age}\times\text{pSES}}\Big).\end{aligned} \tag{4.4}$$

The population coefficient, intercept and two-way interactions of gender$\times$age and pSES$\times$age are assumed to be constant across time and space.

The random effects explored have been a random slope for the seasonal and time trend effect grouped by RHA and a nested random effect of sRHA within RHA. Figure B.8 shows that the coefficient estimates for pSES from the temporal analysis have a large amount of variation between regions. This motivates the inclusion of a random slope for pSES that varies by region. To capture the correlation of regions that are adjacent, a CAR model can

be included. The CAR model is compared to the nested random effects of sRHA within RHA. To capture the temporal correlation an AR(1) model is included. The random effect structures are as follows:

1. An AR(1) model and CAR model random effects: $\boldsymbol{\varepsilon}(t;r) = \psi_t + \eta_r$.
2. An AR(1) model and nested random effects: $\boldsymbol{\varepsilon}(t;r) = \psi_t + u_R + v_r$.
3. An AR(1) model, CAR model and seasonal and time trend slope random effects: $\boldsymbol{\varepsilon}(t;r) = \psi_t + \eta_r + A_{1R}\cos\left(\frac{2\pi t}{13}\right) + B_{1R}\sin\left(\frac{2\pi t}{13}\right) + \phi_{1R}x_{tr\boldsymbol{p}}$.
4. An AR(1) model, CAR model and pSES slope varying by region random effects: $\boldsymbol{\varepsilon}(t;r) = \psi_t + \eta_r + \alpha_{l1r}^{\text{pSES}}$.

Here $\alpha_{\text{treaty},1r}^{\text{pSES}} \sim N(0, \sigma^2_{\alpha_{\text{treaty}}})$, $\alpha_{\text{supported},1r}^{\text{pSES}} \sim N(0, \sigma^2_{\alpha_{\text{supported}}})$, $\boldsymbol{\psi}$ follows an AR(1) model and $\boldsymbol{\eta}$ follows a CAR model. The time trend, seasonal and nested random effects, ϕ_{1R}, A_{1R}, B_{1R}, u_R and v_r, have the same specification used for model ST3. It is assumed that the random slopes A_{1R}, B_{1R} and ϕ_{1R} are correlated and that the levels of the $\alpha_{l1r}^{\text{pSES}}$ random slope, income supported and treaty status, are correlated. We have $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_T) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\psi)$ and $\Sigma_\psi = \sigma^2_\psi/(1-\rho^2)\{\rho^{|t_1-t_2|}\}^T_{t_1,t_2=1}$. The parameter σ^2_ψ is the temporal variance and ρ is the lag-1 temporal autocorrelation. We have $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{70}) \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$ and $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \sigma^2_\eta(\boldsymbol{I} - \lambda_\eta \boldsymbol{W})^{-1}$. Note that in `INLA`, $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$ is estimated as $\boldsymbol{\Sigma}_{\boldsymbol{\eta}} = \sigma^2_\eta\left(\boldsymbol{I} - \frac{\lambda_\eta}{\text{max eigenvalue of } W}\boldsymbol{W}\right)^{-1}$ to ensure λ_η is in the range $[0, 1)$.

Since `INLA` uses Bayesian inference, priors must be assigned to the parameters of the random effects, the hyperparameters. The parameter estimates can be sensitive to the chosen priors (Rue et al., 2009). To let the data speak for itself, we set the priors to be highly vague, but proper. The prior distribution for the AR(1) model parameters are

$$\log\left(\frac{1+\rho}{1-\rho}\right) \sim N(0, 0.15)$$
$$-\log(\sigma^2_\psi/(1-\rho^2)) \sim \log\text{ Gamma}(1, 0.0001),$$

and for the CAR model the prior distributions are

$$-\log(\sigma^2_\eta) \sim \log\text{ Gamma}(1, 0.0001)$$
$$\log\left(\frac{\lambda_\eta}{1-\lambda_\eta}\right) \sim N(0, 10).$$

The prior for ρ is kept at the default in `INLA`. The prior for λ_η is set to be a more vague prior compared to ρ due to the posterior distribution of λ_η being unstable and multimodal at the default $N(0, 0.1)$ recommended by `INLA`. The prior distribution for the random seasonal slope, time trend slope, pSES slope and nested sRHA within RHA effects is set to be the

same highly vague distribution. In other words, we have

$$-\log(\sigma^2) \sim \log \text{Gamma}(1, 0.0001),$$

where σ^2 corresponds to the random seasonal slope, time trend slope, pSES slope, RHA intercept or sRHA intercept variance parameter.

To assess whether the model fit improves with the inclusion of different random effects, the deviance information criteria (DIC) and Watanabe-Akaike information criteria (WAIC) are used. A lower DIC and WAIC compared to other nested models means the model is a better fit. The analysis finds that the inclusion of the random effect u_R, which specifies sRHAs in the same RHA are correlated, does not improve the fit of the model when the random effect v_r is included. In addition, the CAR model random effect does not offer a significantly better fit in the model compared to the model with just a random effect for the sRHAs, v_r. The DIC is 127,694 for the spatio-temporal model with AR(1) model and v_r random effects and 127,693 for the AR(1) and CAR model random effects. This is shown in Table A.11. This could be due to the estimated covariance between sRHAs in the same RHA and the estimated covariance between adjacent sRHAs being small, approximately 0.01. The model could include either a CAR model, nested random effects or a random effect for the sRHA since there is a negligible difference between their information criterions and number of effective parameters.

From Table A.11, we see the model fit improves according to the DIC and WAIC measures when the random slopes for the seasonal and time trend and levels of pSES are included. Since there is a large amount of variability even at the sRHA level, we explore whether the model fit improves by having the random slopes grouped by sRHA, rather than RHA. We see that the model fit sees the most improvement when the pSES slope is grouped by sRHA even with the increased number of effective parameters. The model DIC and WAIC sees a small improvement when the seasonal and time trend slopes are group by sRHA. Overall, the best model is the model with random slopes for pSES and seasonal and time trend grouped by sRHA and random AR(1) and CAR model effects. However, since the DIC and WAIC sees no significant improvement with the CAR model compared to the sRHA random effect, the random effect for the sRHA, v_r, is chosen over the CAR model in favour of a simpler model.

To explore the effect of the prior distribution on the model and parameter estimates, we assign weakly informative prior distributions to the variance parameters. The prior distributions for ρ and λ_η remain the same but the remaining parameters are set such that $-\log(\sigma^2) \sim \log \text{Gamma}(0.1, 0.1)$. The parameter estimates with weakly informative prior distributions are similar aside from the estimate for the RHA random effect and the random slope for the time trend grouped by region. The highly vague prior distributions estimated these effects to be very small but the weakly informative prior distributions estimated them

to be larger due to a stricter distribution setting. Despite the differences in these parameter estimates, the information criterions remain largely the same. The model with only an AR(1) random effect has the largest deviation in information criterions. The information criterions for the models run under weakly informative prior distributions are summarized in Table A.12.

4.3.3 Proposed model given spatio-temporal analysis

The best model according to the WAIC and DIC measures is found to have AR(1) model, sRHA, seasonal slope, time trend slope and pSES slope random effects. It has the following specification,

$$\begin{aligned}
\boldsymbol{\theta}_{\boldsymbol{p}}^{\text{risk/exp}}(t; R^{*,\text{group}}) &= \left(1, \theta_0, \alpha_j^{\text{age}}(t; R^{*,\text{group}}), \alpha_k^{\text{gender}}(t; R^{*,\text{group}}), \alpha_l^{\text{pSES}}, \alpha_{jk}^{\text{age}\times\text{gender}}, \alpha_{jk}^{\text{age}\times\text{pSES}}\right), \\
\boldsymbol{\beta} &= A_0 \cos\left(\frac{2\pi t}{13}\right) + B_0 \sin\left(\frac{2\pi t}{13}\right) + \phi_0 x_{tr\boldsymbol{p}} + \phi_{0l}^{\text{pSES}} x_{tr\boldsymbol{p}}, \\
\boldsymbol{\varepsilon}(t; r) &= \psi_t + v_r + \alpha_{l1r}^{\text{pSES}} + A_{1r} \cos\left(\frac{2\pi t}{13}\right) + B_{1r} \sin\left(\frac{2\pi t}{13}\right) + \phi_{1r} x_{tr\boldsymbol{p}}. \qquad (4.5)
\end{aligned}$$

Note that $\gamma = 1$ as it is fixed as an offset term. A summary of the model's parameters are shown in Table A.13. The summary of this model uses a scaled t, where $t = 1, \ldots, 117$ is scaled to range from $(0, 1]$ by dividing t by its maximum index, 117. This ensures the parameter estimates are on a similar scale where the categorical risk factors are binary with value 0 or 1 and the trigonometric functions sine and cosine range from $[-1, 1]$. All of the credible intervals do not contain zero aside from A_0, B_0 and ϕ_0. This means the seasonal and time trend fixed effects terms are not significant. These fixed effects will remain in the model as the means of their random effects. Originally the model included a quadratic term for time. During the model exploration the 95% credible interval for this estimate included 0 so it was removed from the model. The 95% credible interval for the log of the population size coefficient, γ, contained 1 so it is treated as an offset term in the model. The random slopes of the seasonal and time trend effects were assumed to be correlated. The 95% credible intervals of these correlation estimates all contained 0 and contributed no improvement to the model fit so the random effects are now assumed to be independent of each other. This was also true for the correlation estimates of the random slopes for pSES.

After including an interaction term for gender and age group, our model finds that males aged less than 13 have significantly more MHED visits compared to young females, but females aged greater than 12 have significantly more MHED visits compared to older males. When the individual is aged less than 13, the males from RHA 3 and RPP experience $\exp(0.584) = 1.793$ times more MHED visits compared to females aged less than 13 from RHA 3 and RPP. For males aged greater than 12 from RHA 3 and RPP, they experience $\exp(0.584 - 0.977) = 0.675$ times less visits compared to females. Note how the males aged less than 13 from RHA 9 and RPP pSES group have more visits compared to females, but

on a smaller scale compared to RHA 3, 6 and the other RHAs, $\exp(0.259) = 1.296$. We see that for teenagers, the coefficient estimated by RHA 3 has a smaller magnitude compared to RHA 6, 9 and the other RHAs. The interaction between time and pSES show that the treaty status individuals experience an upward trend in MHED visits, but the income supported individuals experience a downward trend. The estimates of the random effect parameter for the temporal variation, σ^2_ψ, is small, but the temporal autocorrelation estimate, ρ, is fairly large at 0.375. The variance estimates for the seasonal and time trend random slopes are quite small. They could be small due to the AR(1) model random effect capturing the majority of the variation due to time. The estimates for pSES, particularly treaty status, are quite large. The slope of pSES varies between regions.

To check the model fit and whether there are any outliers, the conditional predictive ordinates (CPO) and probability integral transform (PIT) values can be examined. The CPO and PIT are defined as

$$\text{CPO}_i = \pi(y_i^{\text{obs}}|\boldsymbol{y}_{-i}) \qquad \text{and} \qquad \text{PIT}_i = \text{Prob}(y_i^{\text{new}} \leq y_i|\boldsymbol{y}_{-i}). \tag{4.6}$$

Since our response is discrete, the PIT values are modified to: $\text{PIT}_i^* = \text{PIT}_i - 0.5\text{CPO}_i$. The CPO is determined through leave-one-out cross-validation (Rue et al., 2009) and it expresses the posterior probability of observing y_i when the model is fitted to all the data except y_i. Small CPO values suggest that y_i is an outlier and higher values implies a better model fit. A CPO value is considered small when its inverse is greater than 40 and an extreme value when its inverse is greater than 70 (Ntzoufras, 2009). From Figure B.16 it appears there are some outliers, however the amount is relatively small. The small CPO values account for approximately 2% of the data points. The histogram of the CPO values shows that the majority of the values are large. The PIT values are also determined through leave-one-out cross-validation (Rue et al., 2009). For the modified PIT values, if their histogram does not look like a uniform distribution then there may be issues with the model specification. A U-shaped histogram indicates under-dispersed predictive distributions, an inverse U-shaped histogram indicates over-dispersion and a skewed histogram indicates the central tendencies are biased (Czado et al., 2009). In Figure B.16, the histogram of the PIT values shows they are skewed to the left. This indicates that our model tends to have fitted values smaller than the true observations. This could be due to a variety of reasons such as unexplained variation due to variables not included in the model or due to the outliers the CPO values found. Figures B.17 and B.18 show the fitted MHED visits per 1000 sRHA population of the subgroup females aged greater than 12. We see from the plots that the model has difficulty fitting the large MHED rates, like we saw with the PIT values, but the model captures the overall rate.

Chapter 5

Final Remarks

5.1 Summary

In this project, we explore the spatio-temporal pattern of MHED visits made by Albertans aged less than 18. Investigating these patterns could help understand where more resources are needed. This project studies the spatio-temporal patterns by considering the model in two stages, temporal and spatio-temporal. To model the spatio-temporal patterns, we propose a generalized linear mixed effects model. This model identifies important risk factors associated with MHED visits.

In Chapter 2, the descriptive analysis of the MHED visits provides motivation for exploring the spatio-temporal patterns. The figures and numerical summaries of the data show that the rate of MHED visits varies over time and space. The spatial autocorrelation findings show evidence of spatial autocorrelation on the global scale and clustering on the local scale. The MHED visits have a cyclic seasonal pattern according to the temporal autocorrelation analysis.

In Chapter 3 we summarize the proposed modelling separated into two stages. First we consider modelling the temporal patterns using three different models. The first model assigns factors for each of the 13-1 blocks of size 28 days and assumes a linear time trend. The second model uses sine and cosine functions to describe the seasonal pattern and assumes a linear time trend. The third uses a cyclic cubic smoothing spline for the seasonal effect and a thin plate smoothing spline for the time trend. The second stage considers modelling the spatio-temporal patterns by incorporating independent random effects for the sRHA and RHA. The hierarchical structure of the regions is explored by assuming the sRHAs within the same RHA are correlated. The last model of the spatio-temporal modelling stage uses AR(1) and CAR model random effects where the CAR model assumes that adjacent sRHAs are correlated.

In Chapter 4 we conduct the analysis based on the proposed modelling in Chapter 3. The results find that the seasonal effect can be explained by trigonometric functions and that RHAs 3, 6 and 9 risk factor coefficients differ from the other RHAs. We find the effect

of pSES varies over time but the seasonal effect and remaining risk factors/exposures do not. The random effect assuming sRHAs within the same RHA to be correlated is found to have a small estimated covariance. Random slopes for the temporal effects grouped by region improve the model fit according to information criterion measures. The CAR model random effect estimates the covariance between adjacent regions to be small. We find no discernible difference between a model that assumes sRHAs are correlated and one that does not. The large amount of variation over regions for the pSES coefficient motivates the inclusion of a random pSES slope grouped by regions. The model with the best fit according to information criterion measures is found to be the random pSES, seasonal and time trend slopes grouped by sRHA and AR(1) model and sRHA random effects.

5.2 Future Investigation

Our models have shown evidence of spatio-temporal patterns and that demographic groups (gender, age, and socio-economic status) are more susceptible to having MHED visits. There are a variety of issues that could be addressed to further enhance the value of this research. For instance, this data is recurrent event data, meaning that MHED visits made by the same individual are correlated. This project focuses on aggregate-level data but future investigations could incorporate this correlation structure by considering individual-level data. The amount of time elapsed between visits can be studied. The correlation of visits made by the same individual may diminish as the time elapsed between visits increases.

The spatial weights matrix considered in this project is based on the adjacency of two regions. It would be interesting to see how different weight matrices, such as basing the spatial contiguity on a distance threshold, would affect the results. In addition, the spatial and temporal effects are considered separately in our models. A more informative model might include an interaction term for space and time. This can be implemented through a random effect that takes the Kroeneker product of a temporal and spatial structure matrix, such as the AR(1) and CAR structure used in Model ST4.

Our model identified demographic information to be important risk factors. Future investigations plan on including information such as the patient's triage level and diagnosis. This information gives an indication of the severity of the child or youth's MHED visit. Currently, the spatio-temporal Bayesian regression models in the analysis do not explore the residuals of the fitted models. To make the spatio-temporal analysis easily comparable with the other non-Bayesian regression models, we plan on investigating the inclusion of some form of residuals. Lastly, knowing which sRHAs are expected to experience large amounts of MHED visits given their population size, and at what time period to expect these MHED visits could be helpful to the Albertan government. In the future, we plan on investigating the predictive abilities of our model.

Bibliography

Annual Report. Alberta Health Services, 2016-2017.

L. Anselin. Local Indicators of Spatial Association - LISA. *Geographical Analysis*, 27(2): 93–115, 1995.

R. M. Assunçao and E. A. Reis. A new proposal to adjust Moran's I for population density. *Statistics in medicine*, 18:2147–2162, 1999.

Benjamin M. Bolker. Linear and generalized linear mixed models. In Gordon A. Fox, Simoneta Negrete-Yankelevich, and Vinicio J. Sosa, editors, *Ecological Statistics: Contemporary theory and application*, chapter 13. Oxford University Press, 2015.

A. D. Cliff and J. K. Ord. *Spatial Autocorrelation*. Pion, 1973.

C. Czado, T. Gneiting, and L. Held. Predictive Model Assessment for Count Data. *Biometrics*, 65(4):1254–1261, 2009.

Calculating Small Area Analysis: Definition of Sub-region Geographic Units in Alberta. Health Surveillance Branch Alberta Health and Wellness, 2003.

G. F. Jenks. Optimal data classification for choropleth maps. Occasional Paper No. 2, Department of Geography, University of Kansas, Lawrence, 1977.

G. F. Jenks and F. C. Caspall. Error on choroplethic maps: definition, measurement, reduction. *Annals of the association of American geographers*, 61(2):217–244, 1971.

M. Kulldorff and N. Nagarwalla. Spatial disease clusters: Detection and inference. *Statistics in Medicine*, 14(8):799–810, 1995.

K. K. Leitch. *Reaching for the top: A report by the advisor on healthy children and youth*. Health Canada, Ottawa, ON, 2007.

J. LeSage and R. K. Pace. *Introduction to Spatial Econometrics*. CRC Press, 2009.

H. H. Mariathas and R. J. Rosychuk. An examination of three spatial event cluster detection methods. *ISPRS International Journal of Geo-Information*, 4:367–384, 2015.

P. A. P. Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2):17–23, 1950.

A. S. Newton and R. J. Rosychuk. *The Emergency Department Compass: Children's Mental Health. Numerical Companion*. Numerical Companion, Edmonton, AB, 2011.

A. S. Newton, R. J. Rosychuk, S. Ali, D. Cawthorpe, J. Curran, K. Dong, M. Slomp, and L. Urichuk. *The Emergency Department Compass: Children's Mental Health. Pediatric mental health emergencies in Alberta, Canada: Emergency department visits by children and youth aged 0 to 17 years, 2002-2008.* Edmonton, AB, 2011.

A. S. Newton, K. Shave, and R. J. Rosychuk. Does emergency department use for alcohol and other drug use cluster geographically? A population-based retrospective cohort study. *Substance Use and Misuse*, 51(9):1239–1244, 2016.

I. Ntzoufras. *Bayesian Modeling Using WingBUGS.* John Wiley and Sons, 2009.

R. J. Rosychuk, A. S. Newton, X. Niu, and L. Urichuk. Space and time clustering of adolescents' emergency department use and post-visit physician care for mood disorders in Alberta, Canada: A population-based 9-year retrospective study. *Canadian J Public Health*, 106(2):10–16, 2015.

R. J. Rosychuk, D. W. Johnson, L. Urichuk, K. Dong, and A. S. Newton. Does emergency department use and post-visit physician care cluster geographically and temporally for adolescents who self-harm? A population-based 9-year retrospective cohort study from Alberta, Canada. *BMC Psychiatry*, 16:229, 2016.

H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*, 71(2):319–392, 2009.

M. Torabi and R. J. Rosychuk. Spatio-temporal modelling of disease mapping of rates. *Canadian Journal of Statistics*, 38(4):698–715, 2010.

R. Wolfinger and M. O'Connell. Generalized linear mixed models: a pseudo-likelihood approach. *Journal of Statistical Computing and Simulation*, 48(3-4):233–243, 1993.

S. N. Wood. *Generalized Additive Models: An introduction with R.* CRC Press, 2006.

Y. Xiong. Analysis of Spatio-Temporal Data for Forest Fire Control. Master's Project, Department of Statistics and Actuarial Science, Simon Fraser University, 2015.

Appendix A

List of Tables

Table A.1: Number of MHED visits made by children/youth and number of children/youth with an MHED visit summarized by demographic information for each fiscal year.

		Fiscal Year								
		02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11
ED Visits	Total	4,278	4,258	4,472	4629	4,661	4,584	4,849	4,579	4,849
Gender	Female	2,539	2,498	2,612	2,755	2,779	2,667	2,831	2,646	2,833
	Male	1,739	1,760	1,860	1,874	1,882	1,917	2,018	1,933	2,016
Age	0-5	56	62	51	43	53	64	58	70	40
	6-12	437	365	441	478	494	446	467	411	479
	13-18	3,785	3,831	3,980	4,108	4,114	4,074	4,324	4,098	4,330
pSES	RPP	2,734	2,681	2,769	2,789	2,761	2,604	2,877	3,034	3,142
	Treaty Status	521	531	625	657	684	654	754	699	731
	Sponsored	754	795	809	903	929	1,000	858	431	524
	Welfare	269	251	269	280	287	326	360	415	452
Subjects	Total	3,438	3,443	3,643	3,724	3,715	3,663	3,915	3,684	3,773
Gender	Female	2,019	1,981	2,100	2,178	2,153	2,090	2,230	2,100	2,165
	Male	1,419	1,462	1,543	1,546	1,562	1,573	1,685	1,584	1,608
Age	0-5	55	57	49	43	52	64	58	69	40
	6-12	375	318	384	384	393	377	413	365	412
	13-18	3,008	3,068	3,210	3,297	3,270	3,222	3,444	3,250	3,321
pSES	RPP	2,247	2,232	2,306	2,319	2,237	2,140	2,373	2,510	2,552
	Treaty Status	418	431	496	525	534	527	594	562	541
	Sponsored	558	584	620	672	722	739	672	301	338
	Welfare	215	196	221	208	222	257	276	311	342

Table A.2: Number of children/youth, number of MHED visits made by children/youth, number of children/youth who had an MHED visit and the ratio of MHED visits over subjects with an MHED visit for each fiscal year and RHA.

		Fiscal Year								
RHA	Counts	02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11
R1	Population	42,713	42,491	42,288	41,956	42,949	43,859	44,486	45,341	46,256
	ED visits	221	201	221	201	253	250	225	221	241
	Subjects	175	175	192	183	217	194	186	180	194
	ED/Subject	1.26	1.15	1.15	1.10	1.17	1.29	1.21	1.23	1.24
R2	Population	25,883	25,975	26,083	26,251	26,871	27,273	27,491	27,998	28,127
	ED visits	114	124	157	150	139	164	164	148	170
	Subjects	94	103	131	134	128	141	149	124	135
	ED/Subject	1.21	1.20	1.20	1.12	1.09	1.16	1.10	1.19	1.26
R3	Population	280,748	283,019	285,594	291,913	299,405	305,330	313,795	321,110	328,886
	ED visits	1,143	1,173	1,241	1,251	1,363	1,292	1,364	1,385	1,519
	Subjects	933	992	1,042	1,013	1,060	1,040	1,129	1,136	1,215
	ED/Subject	1.23	1.18	1.19	1.23	1.29	1.24	1.21	1.22	1.25
R4	Population	80,012	79,929	79,544	79,711	81,100	81,847	82,320	82,830	83,254
	ED visits	522	499	526	631	517	533	623	579	558
	Subjects	410	404	416	517	430	461	497	474	448
	ED/Subject	1.27	1.24	1.26	1.22	1.20	1.16	1.25	1.22	1.25
R5	Population	29,231	28,482	28,666	28,540	29,235	29,573	29,826	29,787	29,785
	ED visits	131	180	127	130	143	150	167	148	130
	Subjects	105	123	102	108	115	119	123	121	101
	ED/Subject	1.25	1.46	1.25	1.20	1.24	1.26	1.36	1.22	1.29
R6	Population	245,407	245,066	244,019	245,840	252,696	256,368	262,195	266,615	270,921
	ED visits	1,408	1,416	1,359	1,389	1,289	1,273	1,285	1,217	1,339
	Subjects	1,113	1,096	1,091	1,085	1,007	975	1,034	965	1,017
	ED/Subject	1.27	1.29	1.25	1.28	1.28	1.31	1.24	1.26	1.32
R7	Population	53,905	52,257	52,119	51,405	51,606	51,760	51,645	51,733	51,374
	ED visits	375	321	418	434	483	462	462	442	435
	Subjects	306	261	335	331	380	370	373	345	329
	ED/Subject	1.23	1.23	1.25	1.31	1.27	1.25	1.24	1.28	1.32
R8	Population	39,469	39,520	39,447	39,946	40,896	41,499	41,988	42,470	43,148
	ED visits	251	215	259	297	271	301	372	274	311
	Subjects	209	181	206	242	221	244	279	206	226
	ED/Subject	1.20	1.19	1.26	1.23	1.23	1.23	1.33	1.33	1.38
R9	Population	23,089	23,479	23,795	23,845	24,588	25,262	25,757	26,815	27,822
	ED visits	109	126	160	140	194	154	181	158	140
	Subjects	89	105	125	106	151	115	139	127	103
	ED/Subject	1.22	1.20	1.28	1.32	1.28	1.34	1.30	1.24	1.36

Table A.3: Weighted average percentages of children/youth with an MHED visit in Alberta

		Fiscal year								
		02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11
	Overall	0.52	0.52	0.54	0.56	0.55	0.53	0.55	0.51	0.53
Gender	Females	0.31	0.30	0.32	0.33	0.33	0.31	0.32	0.30	0.31
	Males	0.21	0.21	0.23	0.23	0.22	0.22	0.23	0.22	0.22
Age	0-6	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
	7-12	0.05	0.04	0.05	0.06	0.06	0.05	0.05	0.05	0.05
	13-18	0.46	0.47	0.48	0.49	0.48	0.47	0.49	0.46	0.48
pSES	RPP[a]	0.33	0.33	0.34	0.34	0.32	0.30	0.33	0.34	0.34
	Treaty Status	0.06	0.06	0.08	0.08	0.08	0.08	0.09	0.08	0.08
	Sponsored	0.09	0.10	0.10	0.11	0.11	0.12	0.10	0.05	0.06
	Welfare	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.05	0.05

[a]Regular Plan Participant.

Table A.4: Number of classes, goodness of variance fit (GVF) measures, and tabular accuracy index (TAI) for different generalizing techniques on the number of MHED visits per 1000 population by sRHA.

Generalizing technique	Classes	GVF	TAI
Equal intervals	5	0.873	0.693
Quantile breaks	5	0.878	0.617
kmeans	5	0.912	0.682
Fisher-Jenks natural breaks	5	0.923	0.724

Table A.5: Global Moran's I and EBI statistics and their p-values for number of MHED visits. Results are based on 999 random permutations.

	Fiscal year								
	02/03	03/04	04/05	05/06	06/07	07/08	08/09	09/10	10/11
Moran's I	0.063	-0.030	0.044	0.138	0.085	0.129	0.247	0.079	0.028
p-value	0.162	0.554	0.191	0.024	0.085	0.032	0.001	0.120	0.238
EBI	0.079	-0.005	0.063	0.151	0.098	0.145	0.256	0.095	0.036
p-value	0.101	0.418	0.136	0.022	0.054	0.012	0.001	0.071	0.254

Table A.6: Data format of aggregation at the sRHA-level.

r	t	Fiscal Year	Block	Gender	pSES	Age Group
R101	1	2003	1	Male	RPP	<13
R101	2	2003	2	Male	RPP	<13
⋮	⋮	⋮	⋮	⋮	⋮	⋮
R101	117	2011	13	Male	RPP	<13
R101	1	2003	1	Female	RPP	<13
⋮	⋮	⋮	⋮	⋮	⋮	⋮
R101	1	2003	1	Male	Treaty Status	<13
⋮	⋮	⋮	⋮	⋮	⋮	⋮
R101	1	2003	1	Male	Income Supported	<13
⋮	⋮	⋮	⋮	⋮	⋮	⋮
R101	1	2003	1	Male	Income Supported	>12
⋮	⋮	⋮	⋮	⋮	⋮	⋮
R904	117	2011	13	Female	Income Supported	>12

Table A.7: Summary of dispersion parameter estimate for the three proposed temporal models and three coefficient estimation settings.

		Model		
Coefficients varying by	Dispersion parameter	Seasonal factors	Trigonometric	GAM
Constant	estimate	1.47	1.47	1.41
RHA	mean(sd)	1.37(0.10)	1.37(0.10)	1.29(0.09)
	range	1.22 - 1.48	1.20 - 1.50	1.11 - 1.44
sRHA	mean(sd)	1.23(0.22)	1.24(0.23)	1.17(0.08)
	range	0.69 - 1.90	0.72 - 1.96	1.00 - 1.33

Table A.8: Summary of the constant coefficient estimates for the exposure and risk factors from models T1 (seasonal factors), T2 (trigonometric functions) and T3 (smoothing splines), with the standard errors in parentheses.

	Seasonal factors (T1)		Trigonometric (T2)		Splines (T3)	
	Poisson	Neg. Binomial	Poisson	Neg. Binomial	Poisson	Neg. Binomial
Intercept	-9.256 (0.082)	-9.103 (0.097)	-9.263 (0.081)	-9.106 (0.096)	-9.259 (0.081)	-9.094 (0.095)
log(pop.)	0.815 (0.008)	0.795 (0.010)	0.815 (0.008)	0.795 (0.010)	0.815 (0.008)	0.795 (0.010)
Male	-0.352 (0.010)	-0.314 (0.012)	-0.352 (0.010)	-0.313 (0.012)	-0.352 (0.010)	-0.314 (0.012)
Teenage	2.093 (0.016)	2.078 (0.017)	2.093 (0.016)	2.078 (0.017)	2.093 (0.016)	2.078 (0.017)
Treaty	-1.466 (0.015)	-1.432 (0.016)	-1.466 (0.015)	-1.433 (0.016)	-1.466 (0.015)	-1.432 (0.016)
Supported	-0.940 (0.012)	-0.912 (0.014)	-0.940 (0.012)	-0.913 (0.014)	-0.940 (0.012)	-0.913 (0.014)

Table A.9: Summary of the coefficient estimates (± margin of error) for model ST1 and model ST2 with a trigonometric seasonal effect.

	Variable							
Estimate	Intercept	log(S(t;r))	A_0	B_0	Male	Teen	Treaty	Supported
Constant	-9.26(±0.16)	0.82(±0.02)	-0.01(±0.01)	-0.05(±0.01)	-0.35(±0.02)	2.09(±0.03)	-1.47(±0.03)	-0.94(±0.04)
2002/03	-9.82(±0.53)	0.88(±0.06)	0.02(±0.04)	-0.06(±0.04)	-0.38(±0.06)	2.04(±0.10)	-1.66(±0.10)	-0.98(±0.07)
2003/04	-10.32(±0.54)	0.92(±0.06)	0.06(±0.04)	-0.08(±0.04)	-0.35(±0.06)	2.19(±0.10)	-1.62(±0.10)	-0.94(±0.07)
2004/05	-9.53(±0.51)	0.85(±0.05)	-0.05(±0.04)	-0.07(±0.04)	-0.34(±0.06)	2.09(±0.10)	-1.49(±0.09)	-0.94(±0.07)
2005/06	-9.46(±0.49)	0.84(±0.05)	-0.02(±0.04)	-0.07(±0.04)	-0.39(±0.06)	2.07(±0.09)	-1.45(±0.09)	-0.86(±0.07)
2006/07	-8.58±0.47)	0.75(±0.05)	0.03(±0.04)	-0.01(±0.04)	-0.39(±0.06)	2.02(±0.09)	-1.39(±0.09)	-0.82(±0.07)
2007/08	-9.50(±0.48)	0.83(±0.05)	0.02(±0.04)	-0.03(±0.04)	-0.33(±0.06)	2.08(±0.09)	-1.38(±0.09)	-0.67(±0.07)
2008/09	-8.57(±0.45)	0.74(±0.05)	-0.05(±0.04)	-0.06(±0.04)	-0.34(±0.06)	2.11(±0.09)	-1.34(±0.08)	-0.86(±0.07)
2009/10	-8.75(±0.46)	0.76(±0.05)	0.01(±0.04)	-0.02(±0.04)	-0.31(±0.06)	2.14(±0.10)	-1.47(±0.08)	-1.27(±0.08)
2010/11	-9.22(±0.45)	0.81(±0.05)	0.04(±0.04)	-0.07(±0.04)	-0.34(±0.06)	2.12(±0.09)	-1.46(±0.08)	-1.17(±0.07)

Table A.10: AICs and BICs for three different settings of model ST3, a nested generalized linear mixed model.

Seasonal and trend effect	AIC	BIC
Constant: A_0, B_0 and ϕ_0	129,563	129,838
Fixed by RHA: $A_0(t;R^*)$, $B_0(t;R^*)$ and $\phi_0(t;R^*)$	129,555	129,849
Random by RHA: $(A_0 + A_{1R})$, $(B_0 + B_{1R})$ and $(\phi_0 + \phi_{1R})$	129,525	129,829

Table A.11: WAIC and DICs for model ST4, a generalized linear mixed model, with highly vague priors.

Parameters	DIC	Eff. parameters	WAIC	Eff. parameters
ψ_t	131,197	119	131,244	165
ψ_t and u_R	130,627	126	130,675	174
ψ_t and v_r	127,694	188	127,761	253
ψ_t, v_r and u_R	127,693	186	127,760	253
ψ_t and η_r	127,693	186	127,760	252
ψ_t, η_r and u_R	127,693	186	127,760	252
ψ_t, η_r, A_{1R}, B_{1R} and ϕ_{1R}	127,569	201	127,641	272
ψ_t, η_r, A_{1r}, B_{1r} and ϕ_{1r}	127,489	293	127,592	393
ψ_t, η_r, $\alpha_{l1R}^{\text{pSES}}$	126,515	196	126,582	261
ψ_t, η_r, $\alpha_{l1r}^{\text{pSES}}$	119,756	310	119,820	370
ψ_t, η_r, A_{1r}, B_{1r}, ϕ_{1r} and $\alpha_{l1r}^{\text{pSES}}$	119,563	415	119,649	493
ψ_t, v_r, A_{1r}, B_{1r}, ϕ_{1r} and $\alpha_{l1r}^{\text{pSES}}$	119,564	415	119,649	493

Table A.12: WAIC and DICs for model ST4, a generalized linear mixed model, with weakly informative priors.

Parameters	DIC	Eff. parameters	WAIC	Eff. parameters
ψ_t	131,062	121	131,110	169
ψ_t and u_R	130,627	128	130,676	176
ψ_t and v_r	127,693	188	127,761	255
ψ_t, v_r and u_R	127,692	188	127,760	255
ψ_t and η_r	127,692	188	127,760	254
ψ_t, η_r and u_R	127,692	188	127,760	254
ψ_t, η_r, A_{1R}, B_{1R} and ϕ_{1R}	127,573	211	127,649	285
ψ_t, η_r, A_{1r}, B_{1r} and ϕ_{1r}	127,493	335	127,598	399
ψ_t, η_r, $\alpha_{l1R}^{\text{pSES}}$	126,513	197	126,581	260
ψ_t, η_r, $\alpha_{l1r}^{\text{pSES}}$	119,755	312	119,820	371
ψ_t, η_r, A_{1r}, B_{1r}, ϕ_{1r} and $\alpha_{l1r}^{\text{pSES}}$	119,587	469	119,685	557
ψ_t, v_r, A_{1r}, B_{1r}, ϕ_{1r} and $\alpha_{l1r}^{\text{pSES}}$	119,588	469	119,685	558

Table A.13: Summary of model ST4 with scaled t and AR(1), sRHA, pSES, seasonal and time trend random effects.

Variable	Mean	95% CI
Intercept	-11.579	(-11.702, -11.457)
A_0	0.013	(-0.035, 0.062)
B_0	-0.051	(-0.101, -0.001)
t	0.024	(-0.120, 0.167)
Male and RHA 3	0.584	(0.517, 0.650)
Male and RHA 6	0.539	(0.471, 0.608)
Male and RHA 9	0.259	(0.133, 0.383)
Male and Other RHAs	0.467	(0.401, 0.533)
Teenage and RHA 3	2.383	(2.315, 2.451)
Teenage and RHA 6	2.883	(2.807, 2.961)
Teenage and RHA 9	2.683	(2.513, 2.858)
Teenage and Other RHAs	2.712	(2.643, 2.781)
Treaty Status	-2.687	(-3.069, -2.310)
Income Supported	-0.671	(-0.830, -0.511)
Treaty Status×t	0.307	(0.199, 0.415)
Income Supported×t	-0.150	(-0.232, -0.067)
Treaty Status×Teenage	0.274	(0.169, 0.380)
Income Supported×Teenage	-0.357	(-0.427, -0.287)
Male×Teenage	-0.977	(-1.041, -0.913)
σ^2_v	0.108	(0.076, 0.164)
σ^2_ψ	0.017	(0.013, 0.025)
ρ	0.375	(0.194, 0.556)
σ^2_A	0.001	(0.001, 0.004)
σ^2_B	0.003	(0.001, 0.006)
σ^2_ϕ	0.050	(0.031, 0.084)
$\sigma^2_{\alpha_{\text{treaty}}}$	2.223	(1.610, 3.270)
$\sigma^2_{\alpha_{\text{supported}}}$	0.338	(0.246, 0.505)

Appendix B

List of Figures

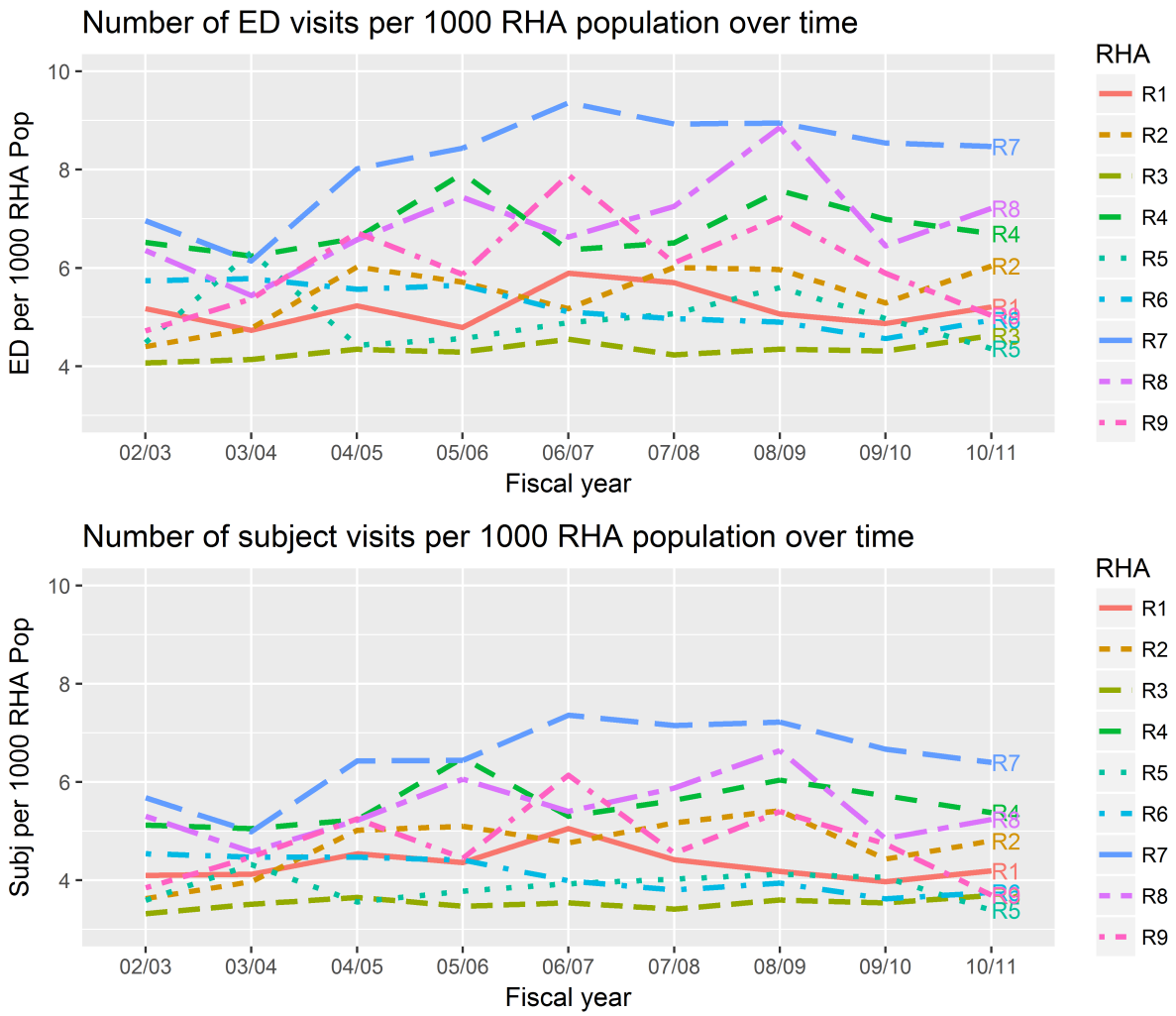

Figure B.1: Plots of MHED visit counts and subject visit counts by RHA over time (yearly).

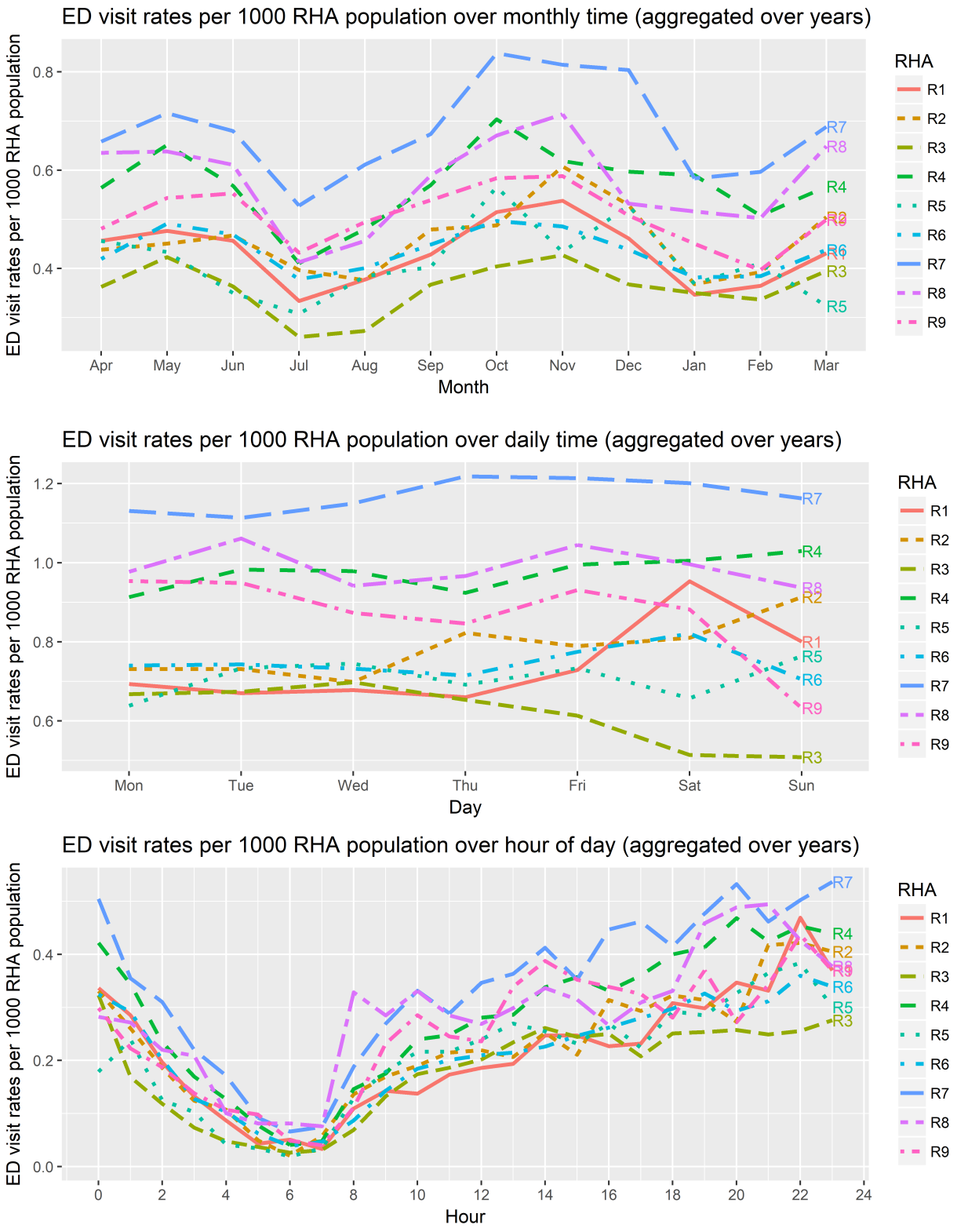

Figure B.2: Plots of MHED visits per 1000 RHA population over months (top), days (middle), and hour of day (bottom) aggregated over all fiscal years.

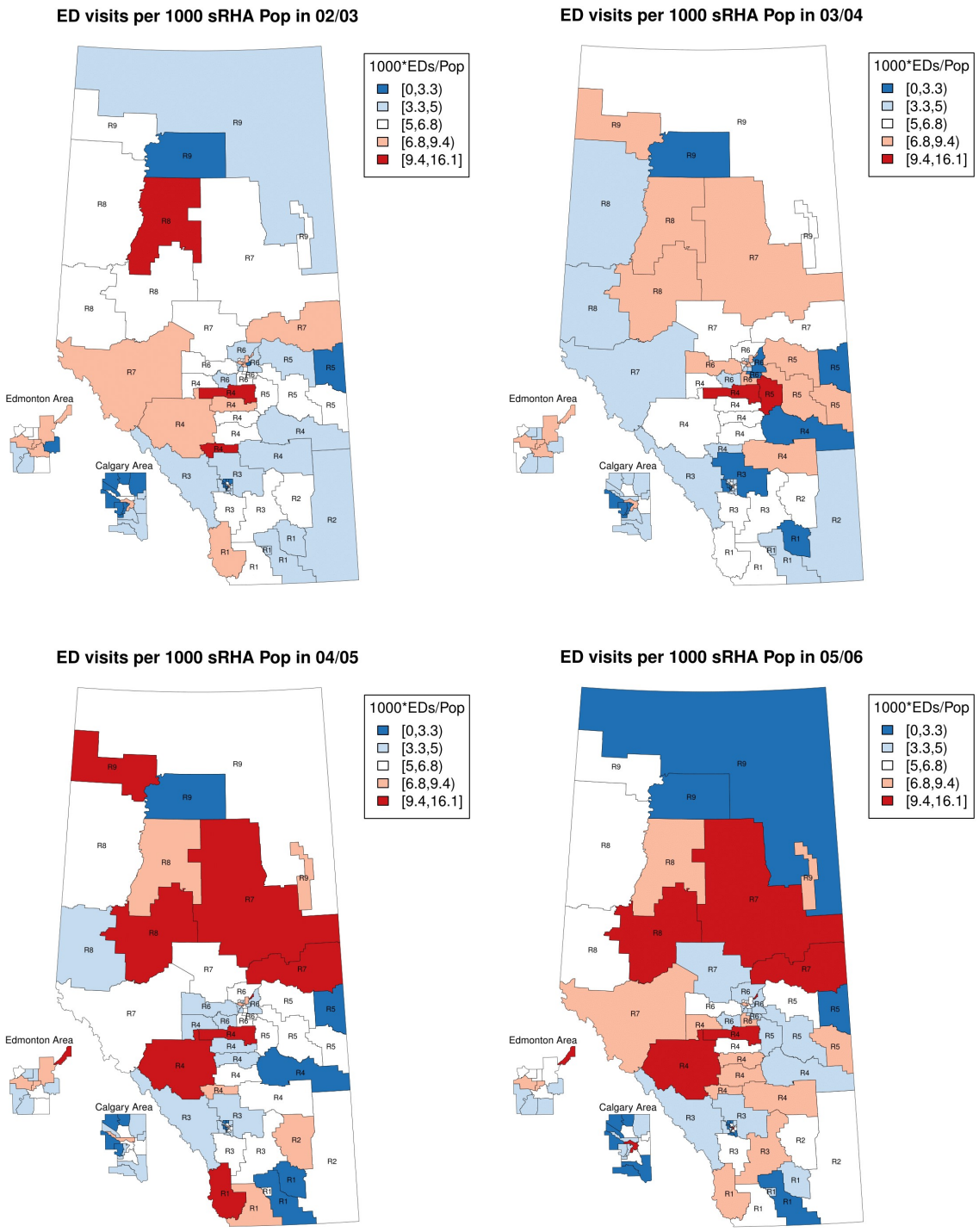

Figure B.3: Choropleth maps of number of MHED visits per 1000 sRHA population over 9 fiscal years.

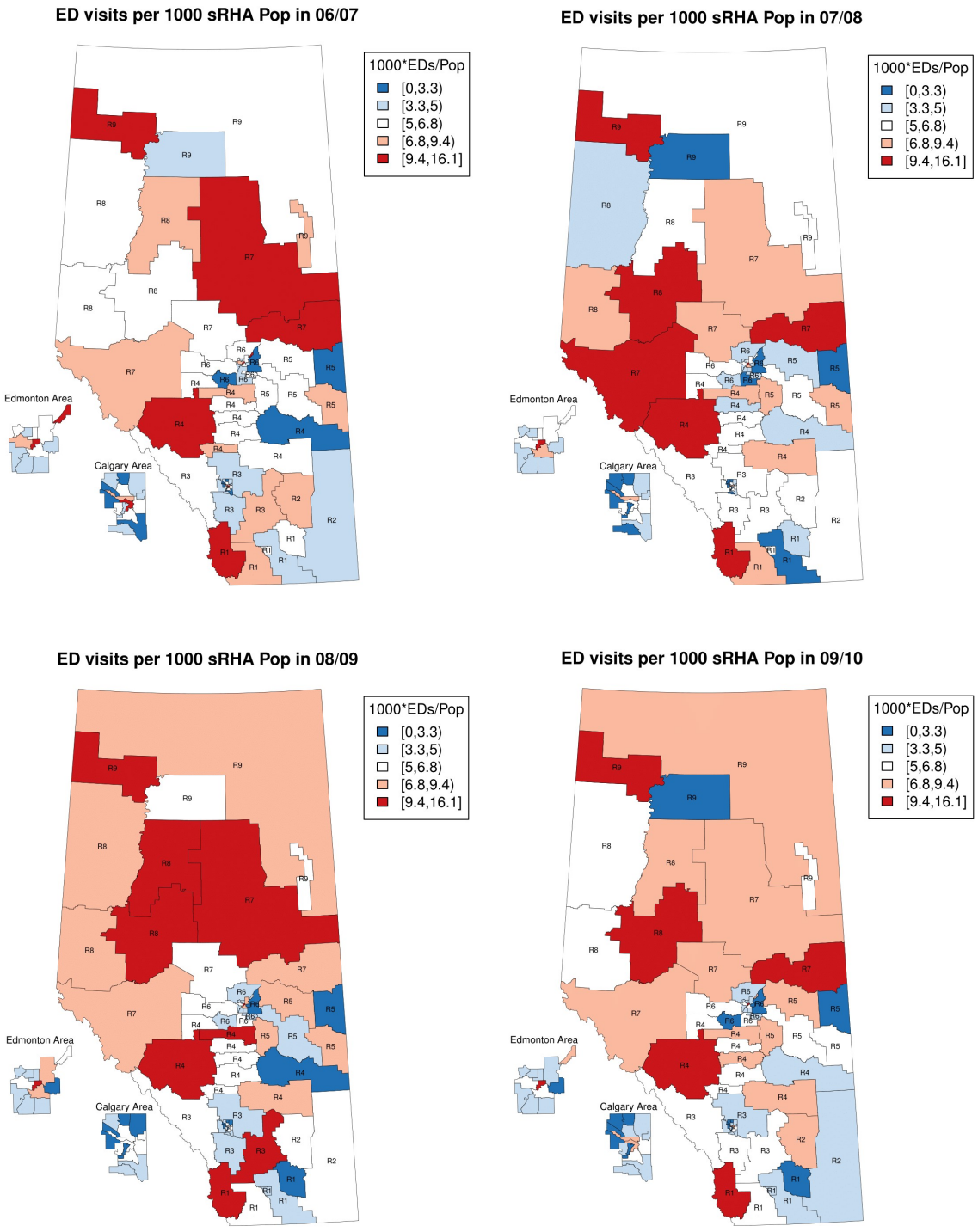

Figure B.3: Choropleth maps of number of ED visits per 1000 sRHA population over 9 fiscal years.

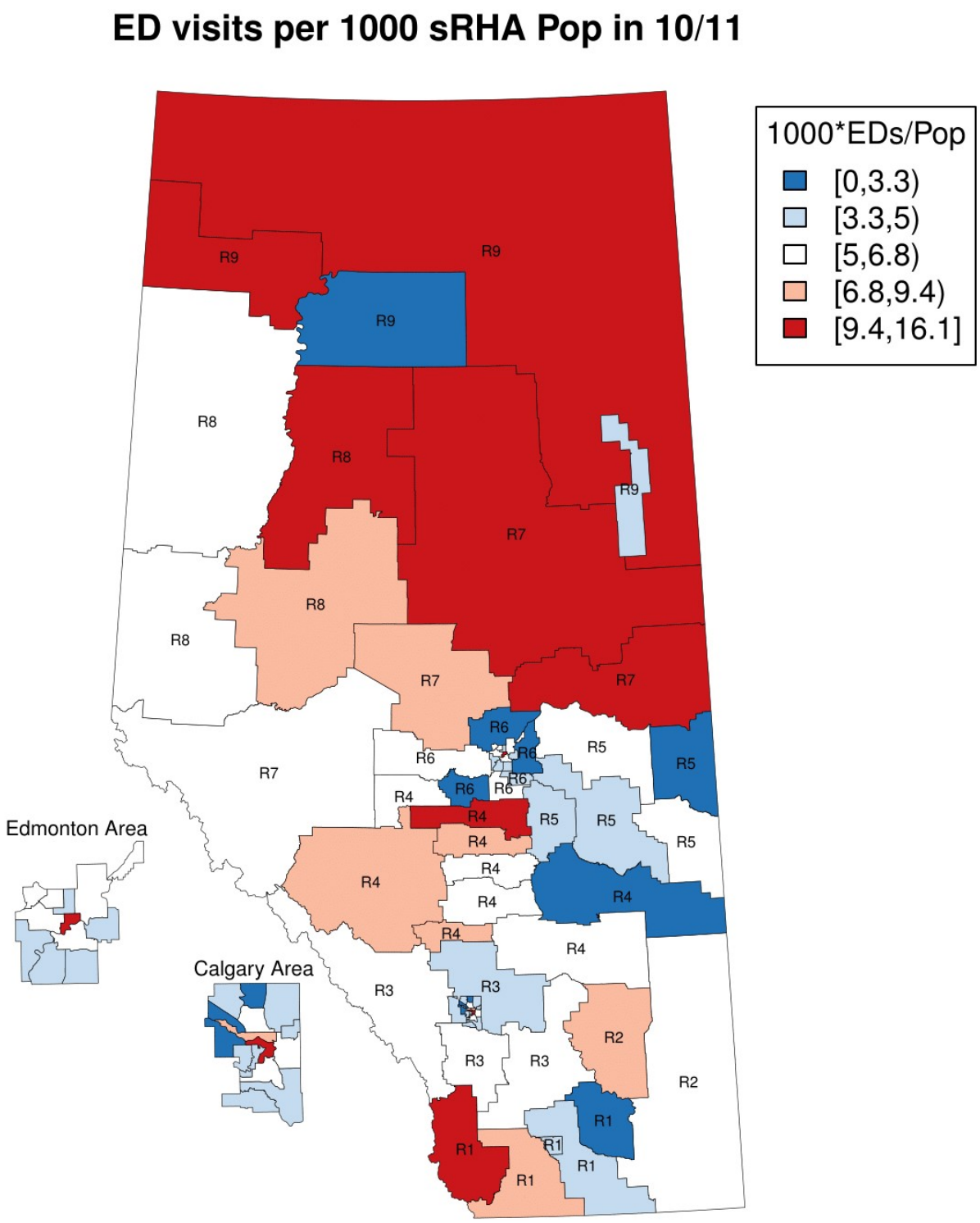

Figure B.3: Choropleth maps of number of MHED visits per 1000 sRHA population over 9 fiscal years.

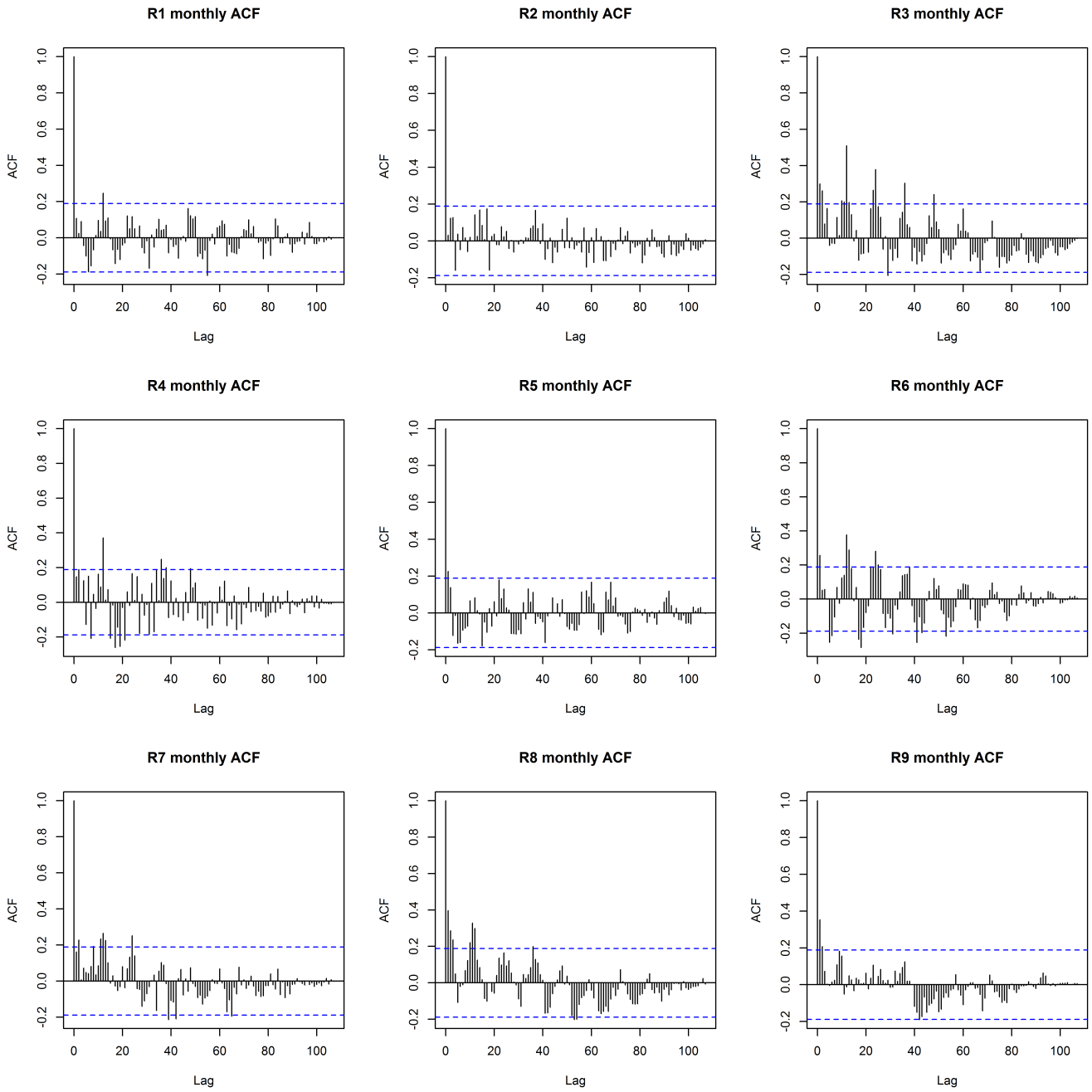

Figure B.4: ACF plots of monthly MHED visits over nine fiscal years for each RHA.

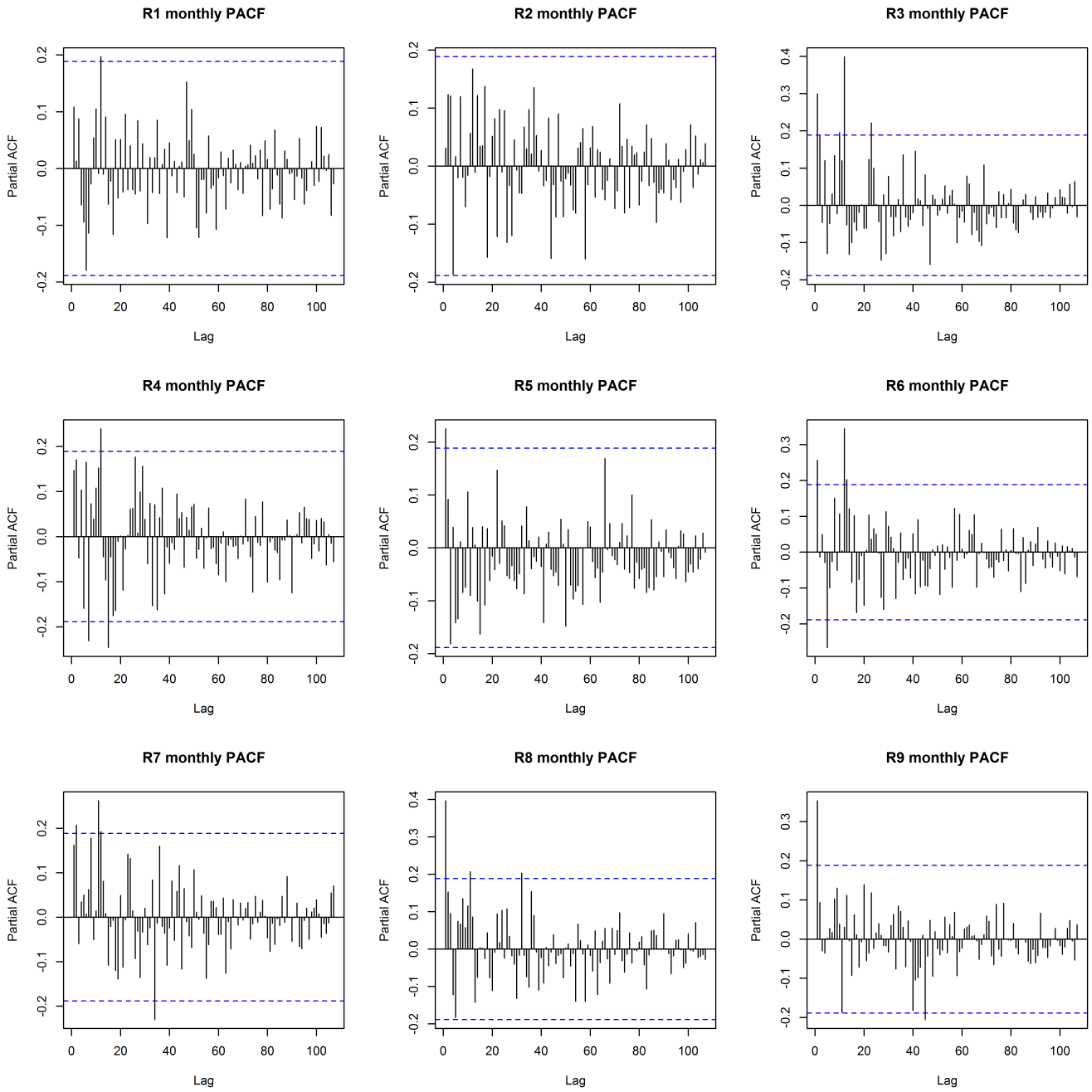

Figure B.5: PACF plots of monthly MHED visits over nine fiscal years for each RHA.

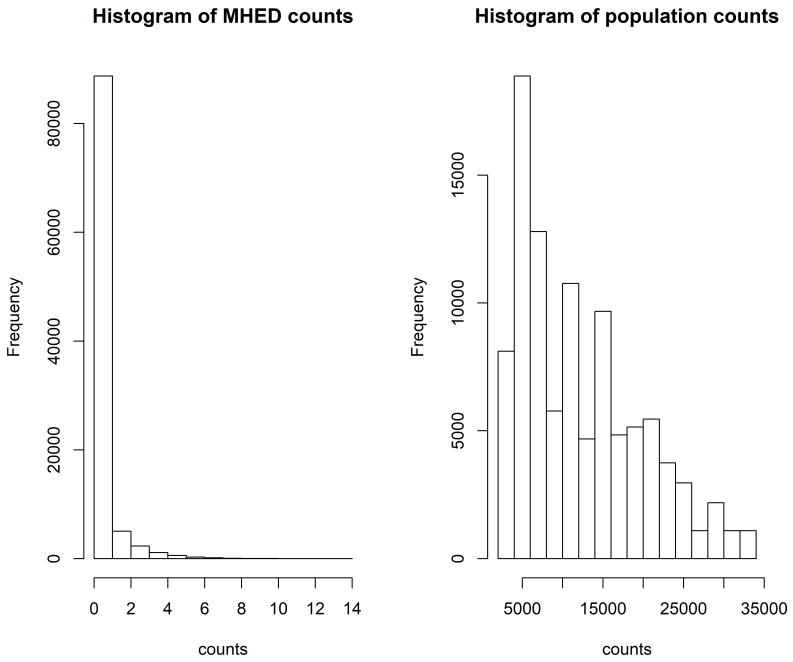

Figure B.6: Distribution of the counts for the number of MHED visits (left) and population counts (right). The MHED visits are stratified by fiscal year, 28 day block, sRHA, gender, age groups 0-12 and 13-18 and pSES. The population counts are stratified by fiscal year and sRHA.

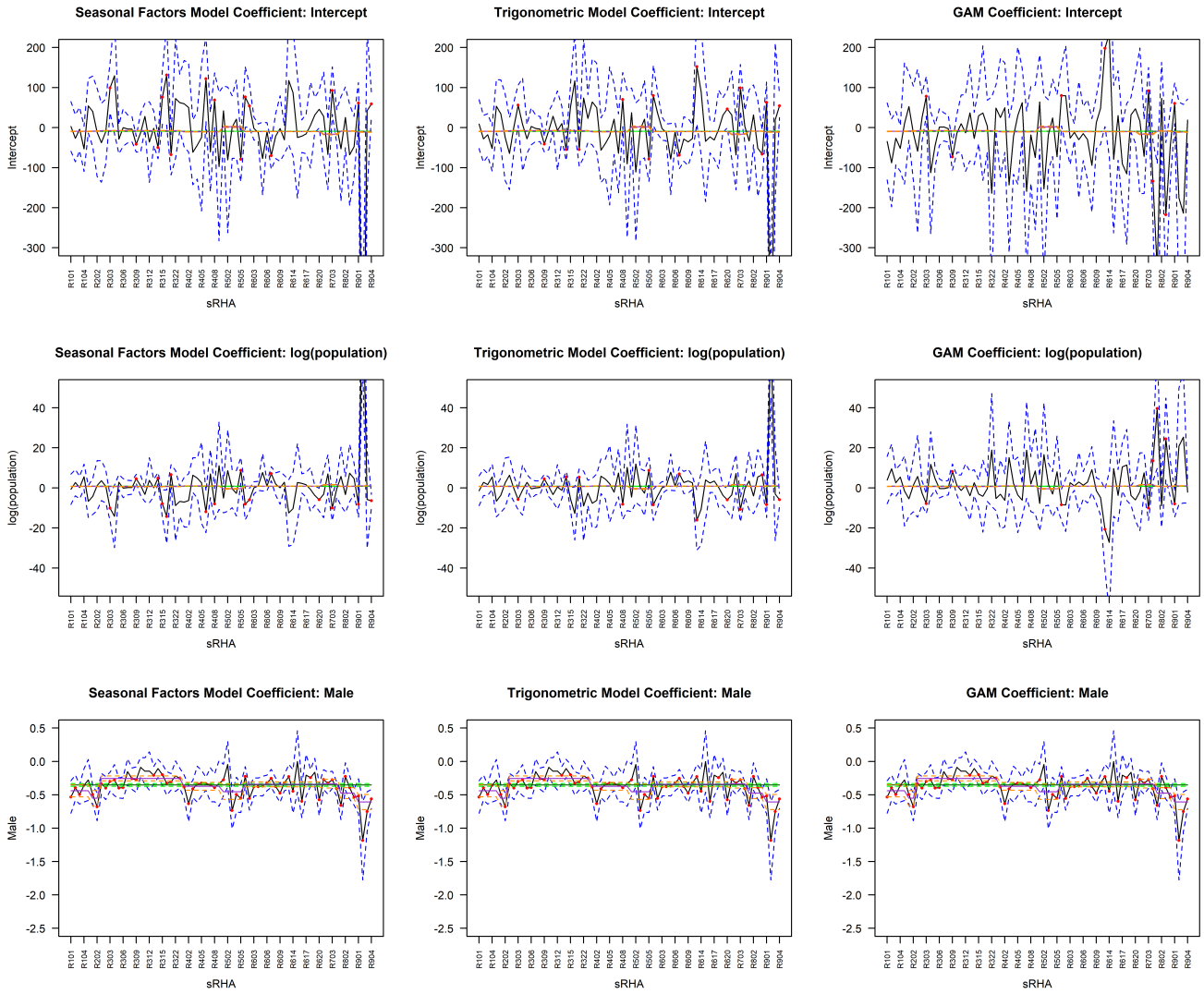

Figure B.7: Part I of the plots of the coefficient estimates for the seasonal factors (left), trigonometric (middle) and smoothing splines (right) models. The blue dotted lines are the 95% CI of the coefficients estimated by sRHA. The black line is the coefficient estimated by sRHA. The red points are the coefficient estimates significantly different from zero. The orange dashed lines are the 95% CI of the coefficients estimated by RHA. The purple line is the coefficient estimate by RHA. The green dashed lines are the 95% CI of the coefficient kept constant across regions.

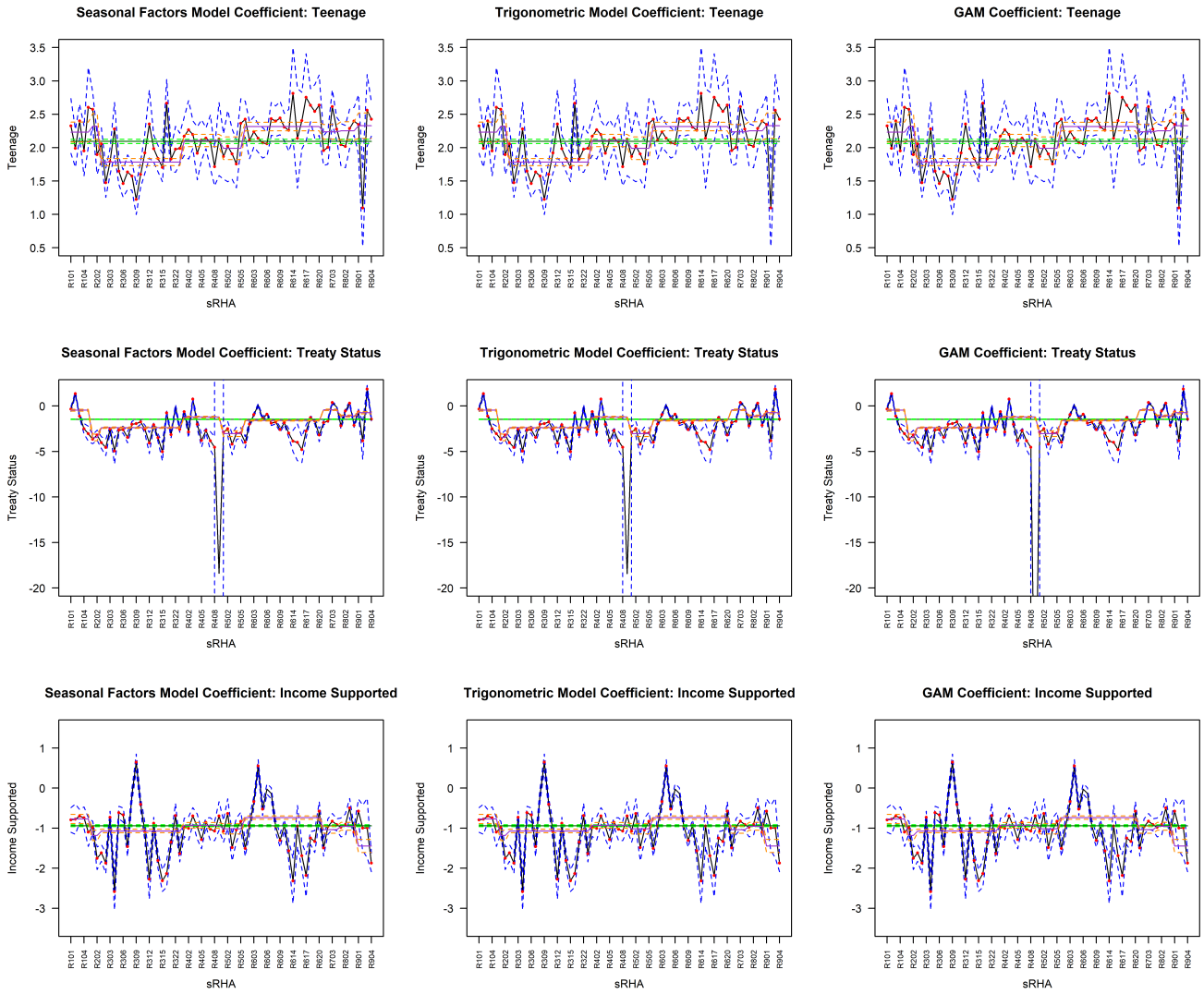

Figure B.8: Part II of the plots of the coefficient estimates for the seasonal factors (left), trigonometric (middle) and smoothing splines (right) models. The blue dotted lines are the 95% CI of the coefficients estimated by sRHA. The black line is the coefficient estimated by sRHA. The red points are the coefficient estimates significantly different from zero. The orange dashed lines are the 95% CI of the coefficients estimated by RHA. The purple line is the coefficient estimate by RHA. The green dashed lines are the 95% CI of the coefficient kept constant across regions.

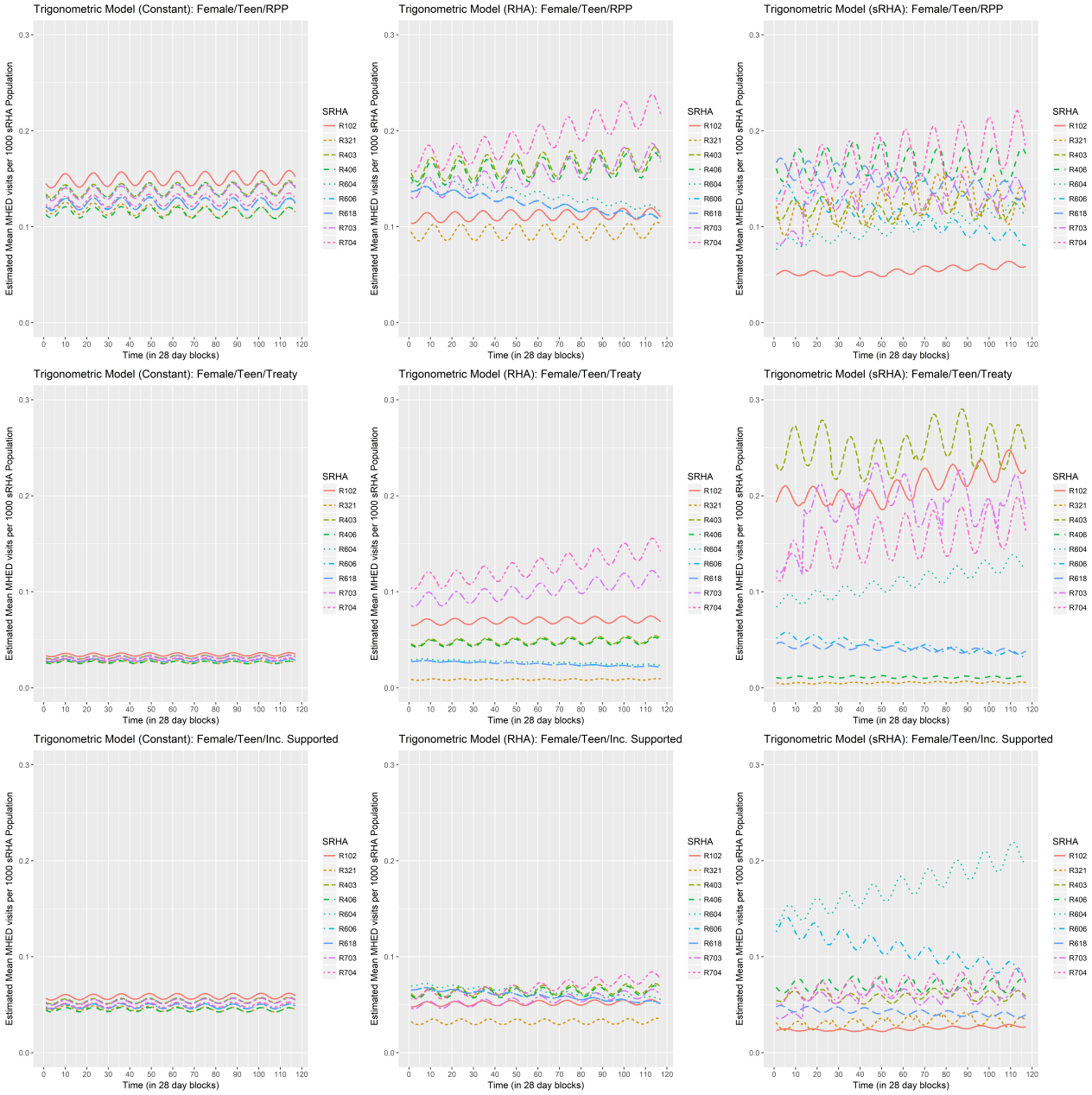

Figure B.9: Plots of the estimated average MHED counts per 1000 sRHA population for select sRHAs for combinations of females, teenagers and the three pSES categories. The plots are the estimated average counts from the trigonometric model T2 where the coefficients were constant (left), estimated by RHA (middle) and estimated by sRHA (right).

Figure B.10: Plots of the estimated average MHED counts per 1000 sRHA population for select sRHAs for combinations of females, teenagers and the three pSES categories. The plots are for the seasonal factors model T1 (left), smoothing splines model T3 (middle) and trigonometric model T2 (right) where the coefficients were estimated by sRHA.

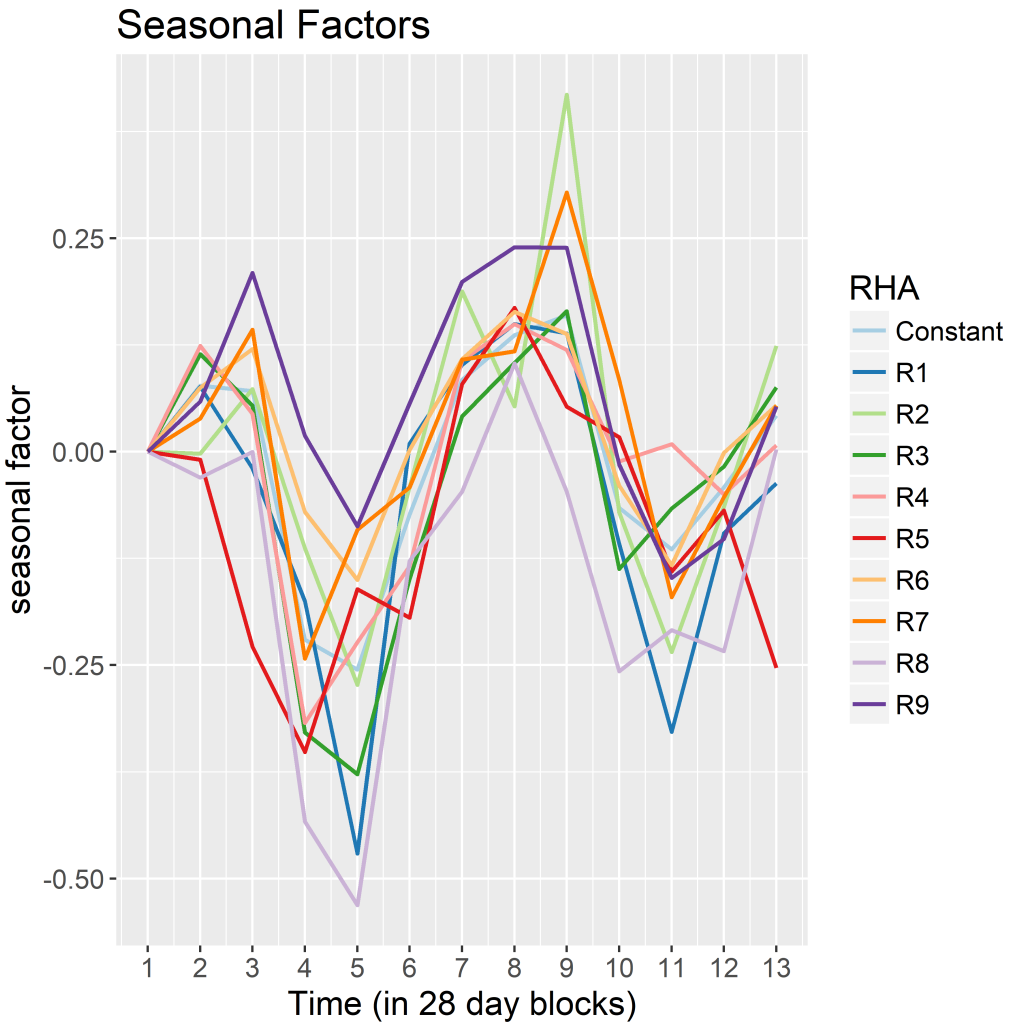

Figure B.11: Plot of the seasonal coefficients for the seasonal factors model.

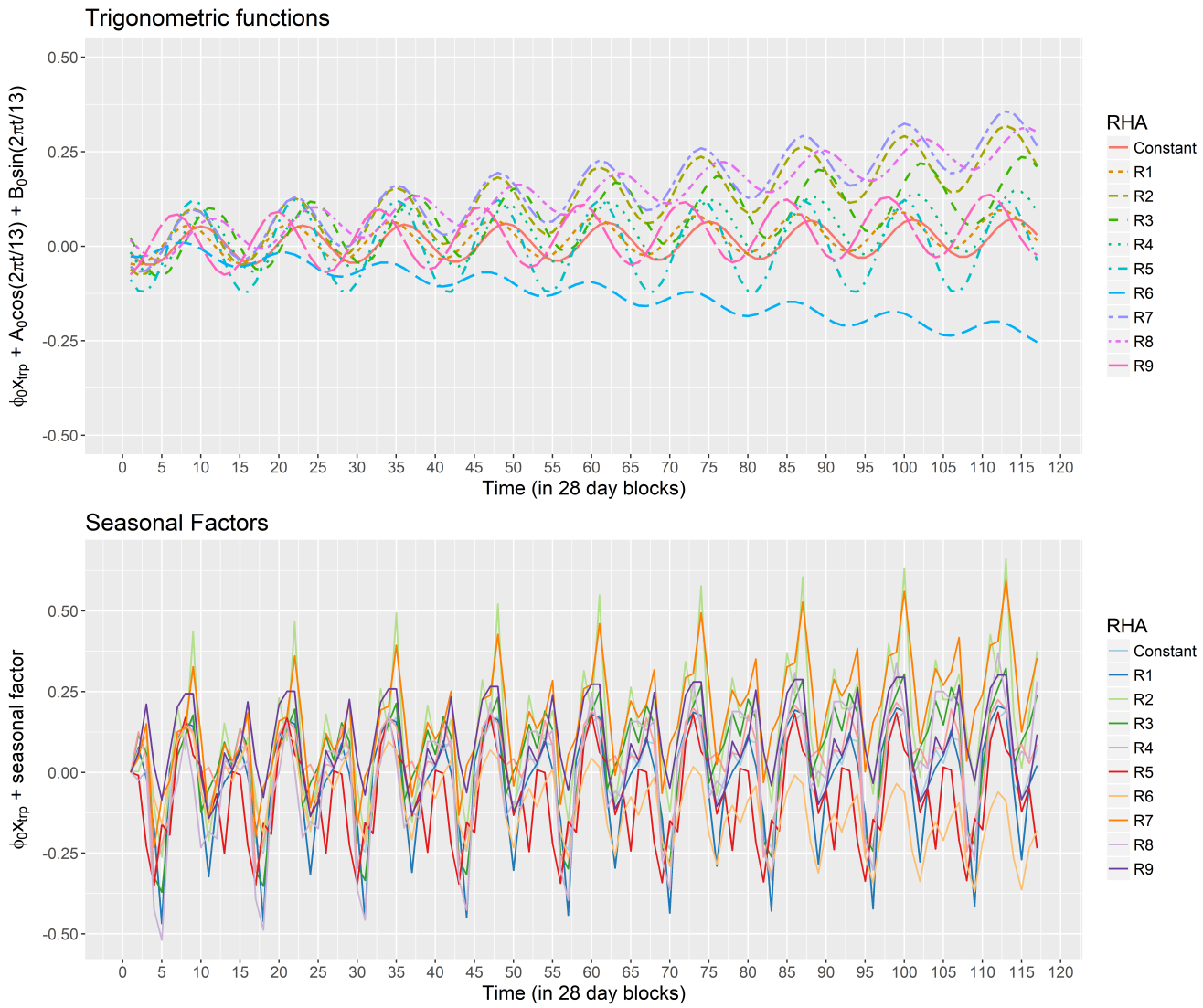

Figure B.12: Plots of the seasonal coefficients and time coefficient over the time unit (28 days) for the seasonal factors and trigonometric models.

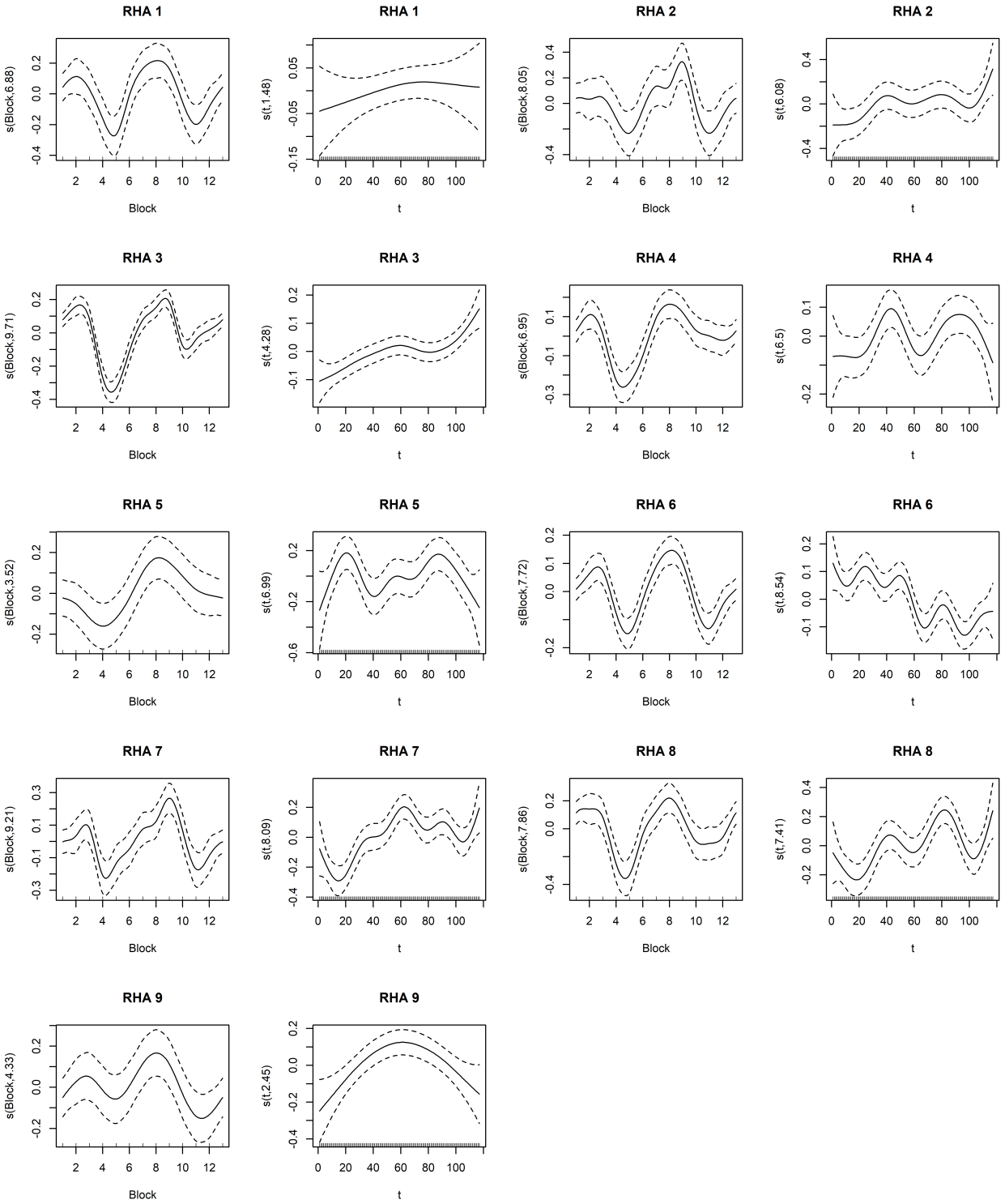

Figure B.13: Plots of the smoothing splines for the blocks and time by RHA from model T3.

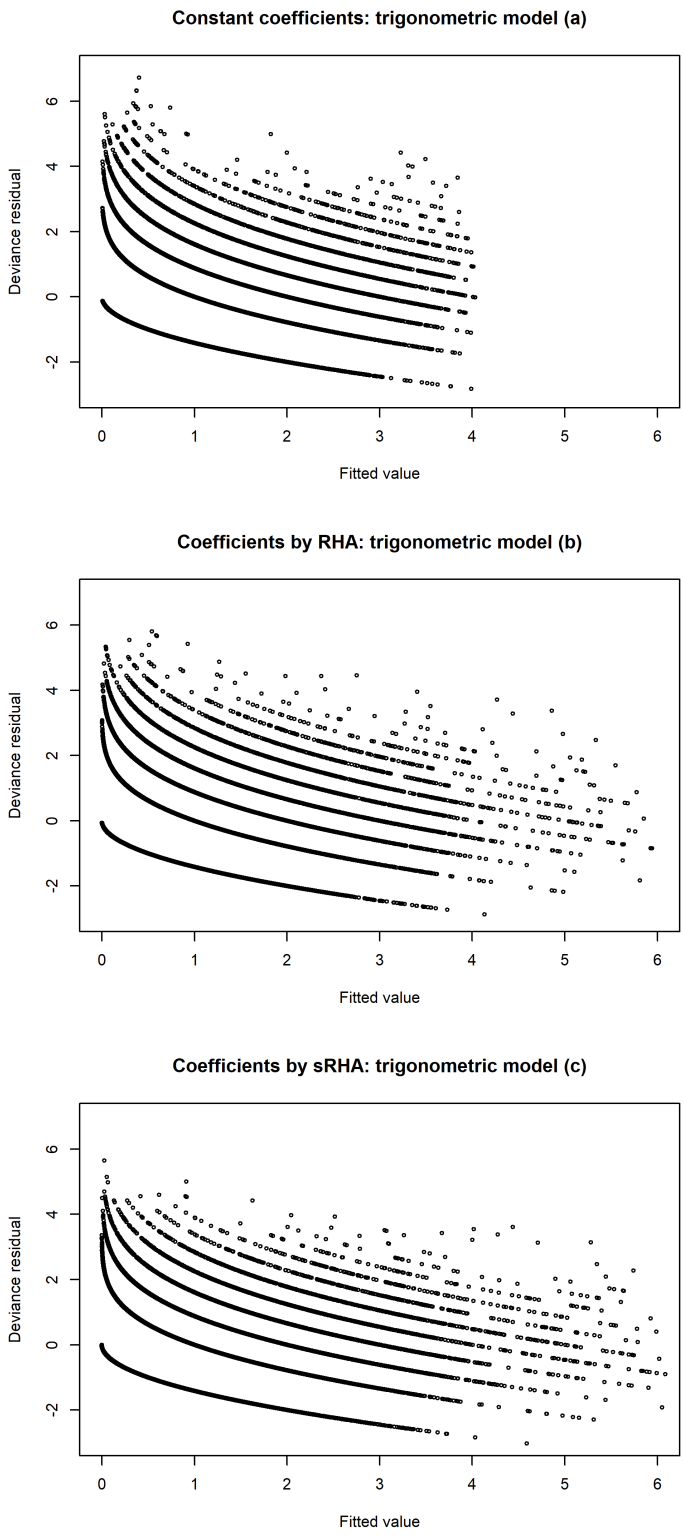

Figure B.14: Deviance residuals vs fitted values of the trigonometric functions temporal model under the three estimation settings.

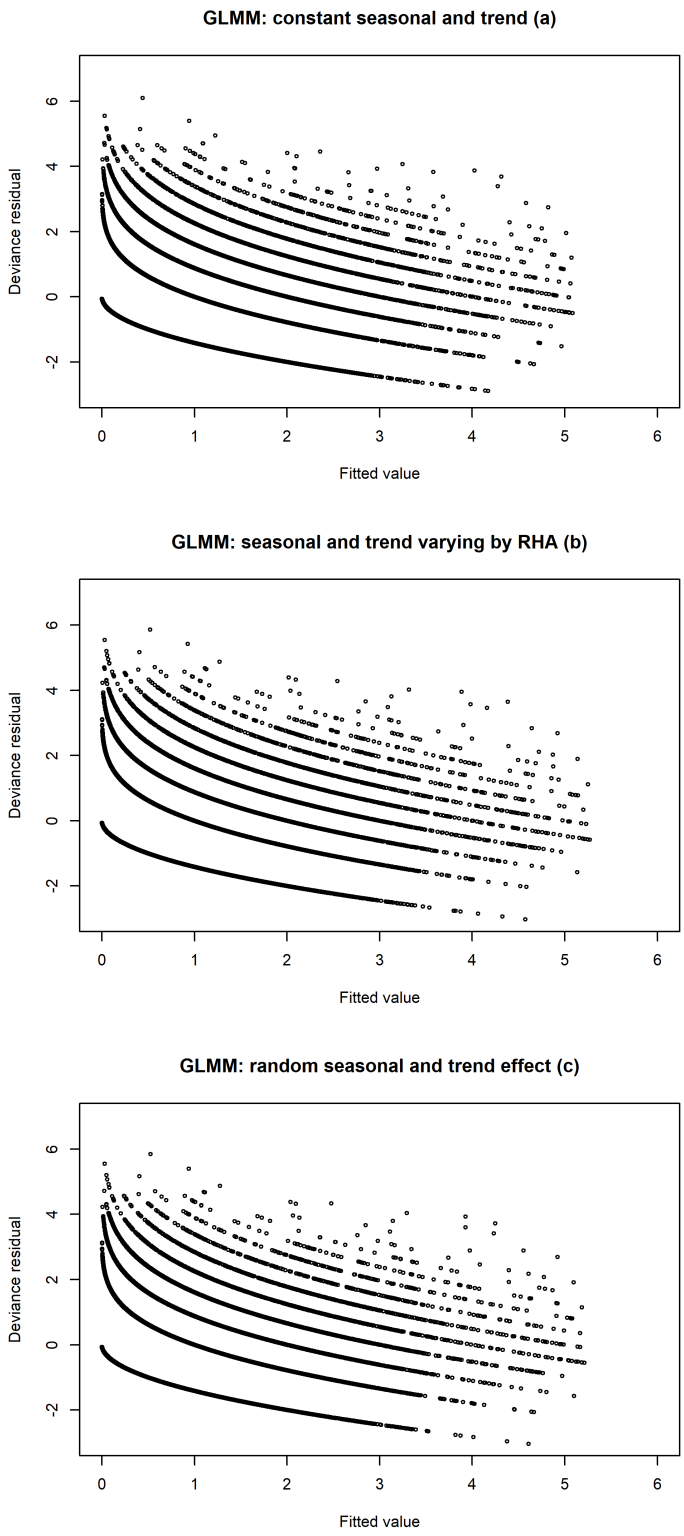

Figure B.15: Deviance residuals vs fitted values of the generalized linear mixed model with nested random effects.

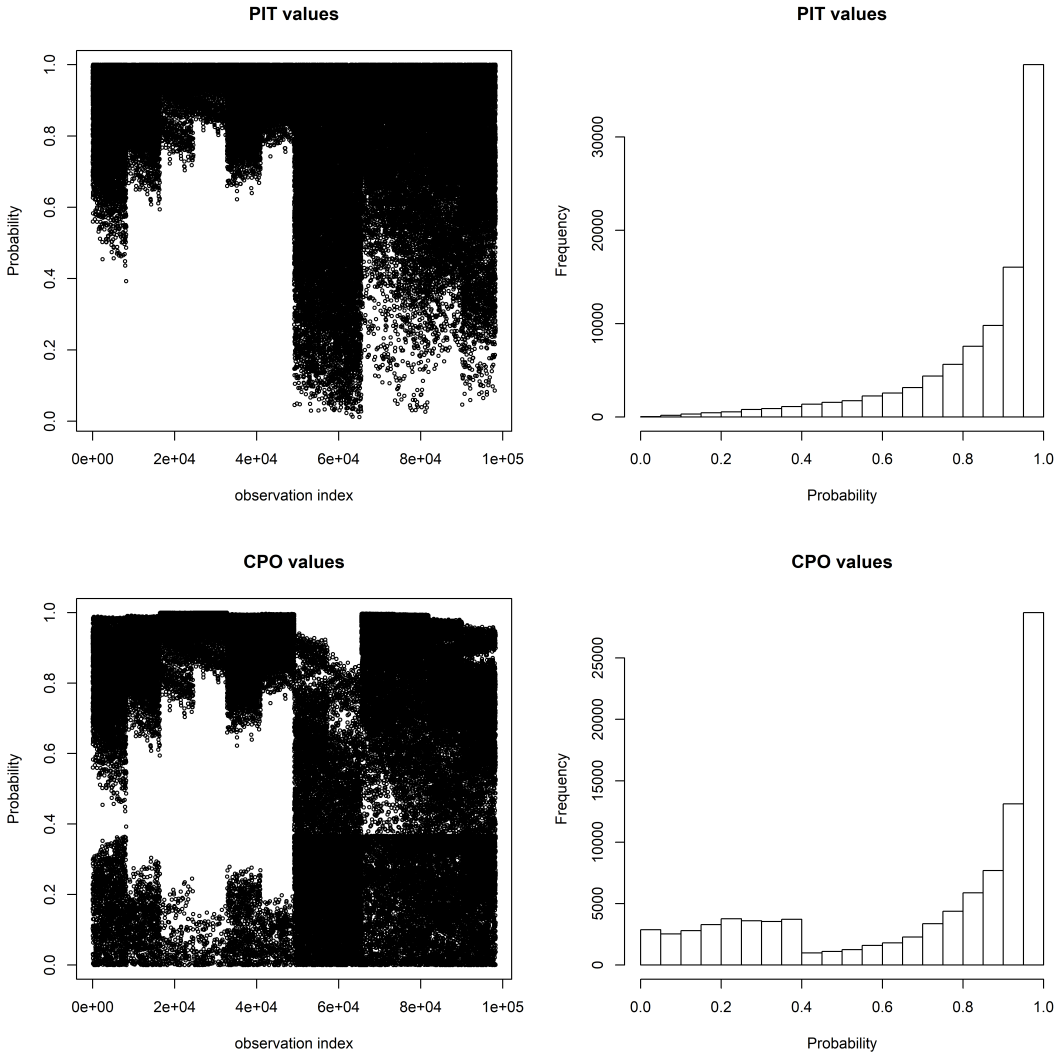

Figure B.16: CPO and PIT values for Model ST4: AR(1), CAR, pSES and seasonal and time trend random effects.

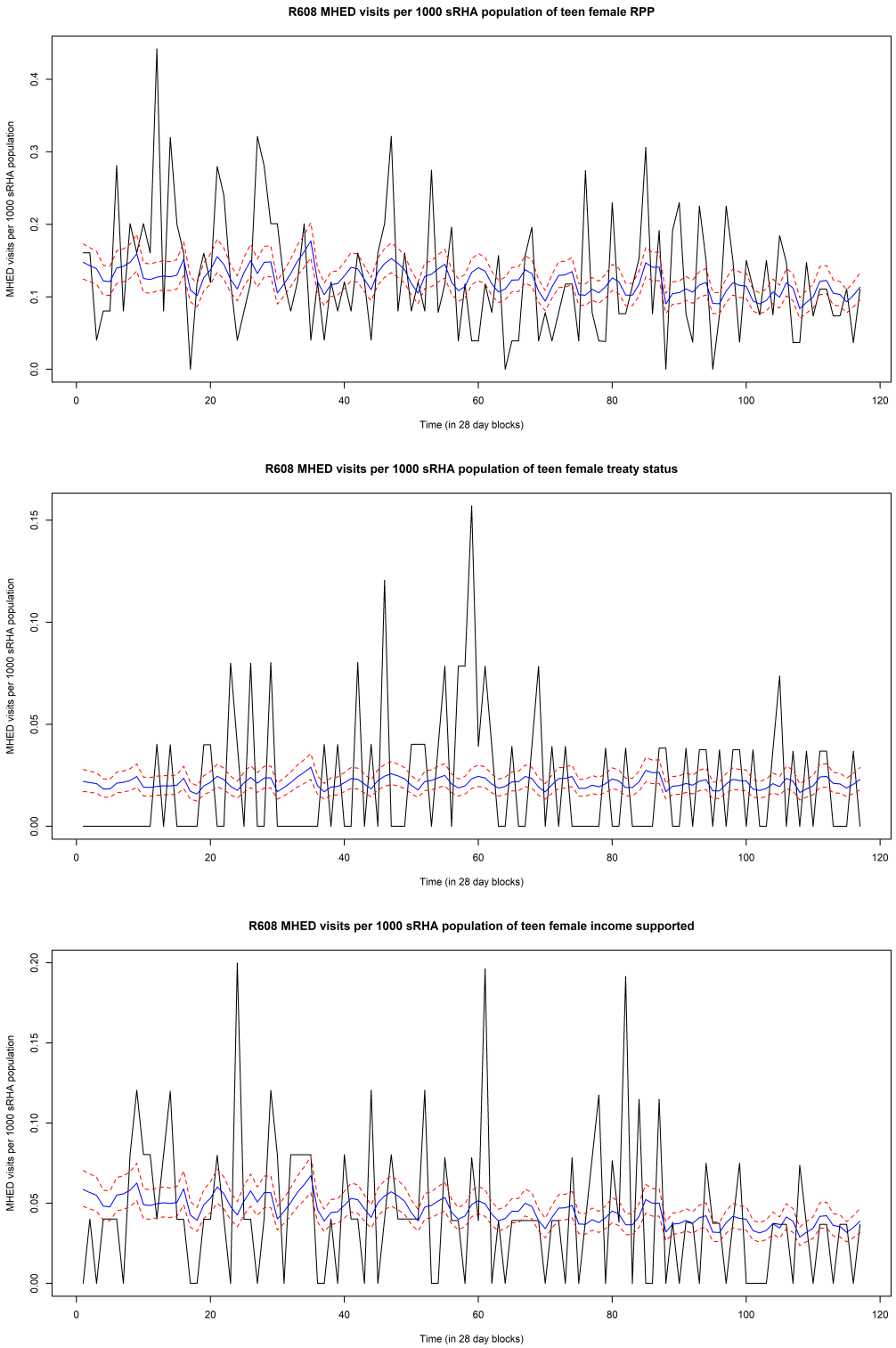

Figure B.17: Plots of MHED visits per 1000 sRHA population for females aged greater than 12 in sRHA R608 who are regular plan participants (RPP), treaty status or income supported. The black line is the true rates, the blue line is the fitted posterior means from Model ST4: AR(1), CAR, seasonal and time trend and pSES random effects. The dashed red lines are the 2.5% and 97.5% quantiles of the fitted posterior mean.

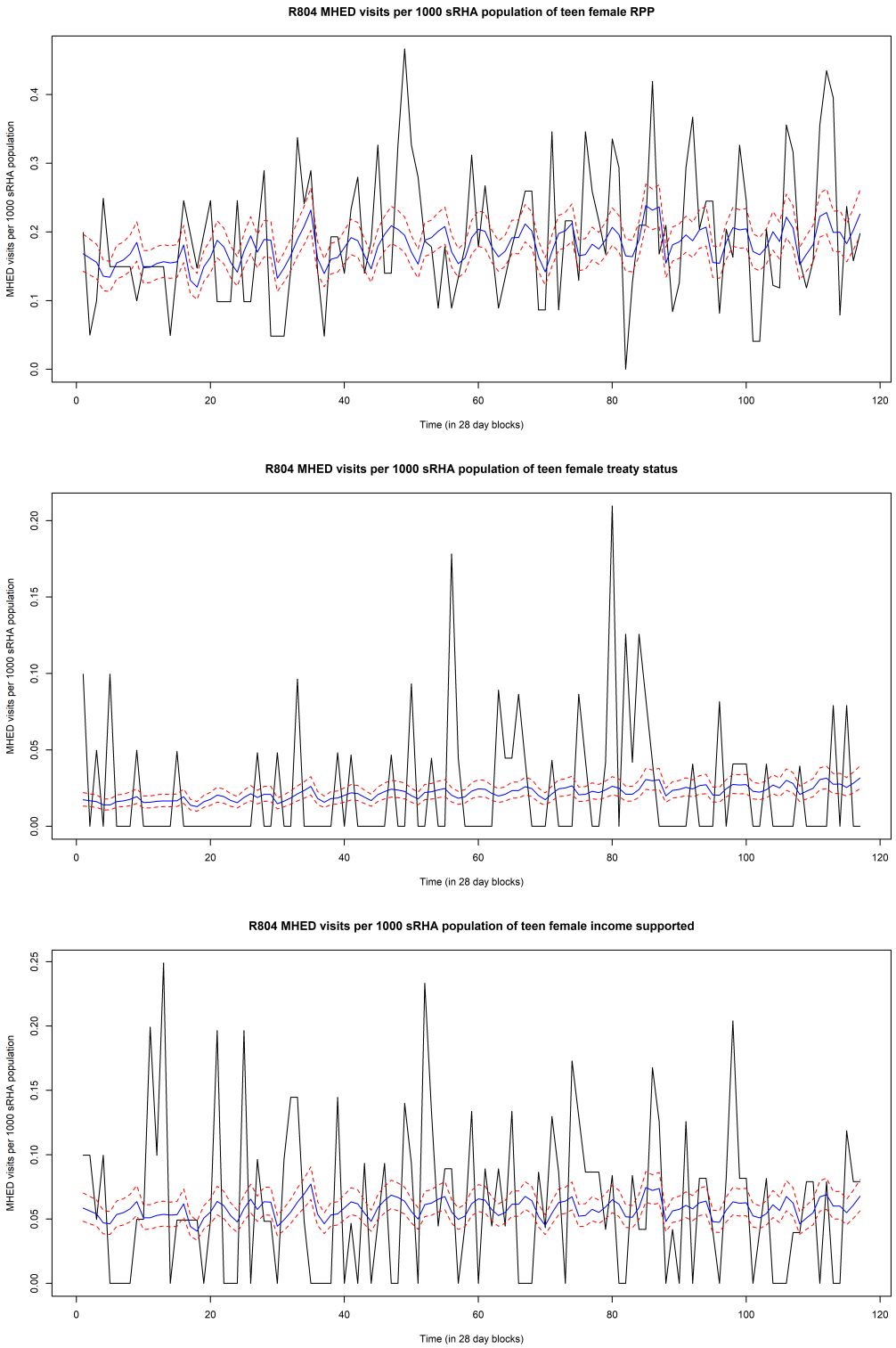

Figure B.18: Plots of MHED visits per 1000 sRHA population for females aged greater than 12 in sRHA R804 who are regular plan participants (RPP), treaty status or income supported. The black line is the true rates, the blue line is the fitted posterior means from Model ST4: AR(1), CAR, seasonal and time trend and pSES random effects. The dashed red lines are the 2.5% and 97.5% quantiles of the fitted posterior mean.