Spatio-temporal Modelling and Mapping of Teenage Birth Rates

Paramjit Gill
Department of Mathematics and Statistics
Okanagan University College, Kelowna, BC, Canada V1V 1V7
pgill@ouc.bc.ca

Abstract

We use generalized linear models to study the effect of availability of family planning services, socioeconomic and demographic factors on the teenage birth rates in the state of Texas. In addition to these factors, we compare the recent downward trend among the white, black and Hispanic teen birth rates. Fitting the generalized linear mixed models using the WinBUGS software incorporates over-dispersion and spatial variation.

The main purpose of our analysis is to measure the magnitude of variation and to assess the role of contributing factors at the geographical level. Key analytical goals are to provide county-specific estimates of the birth rates in the 1990's and to derive estimates of covariate effects, such as race, poverty level, availability of family planning services, degree of urbanity etc. Birth counts data for the 15-17 years old girls from the 254 counties in Texas over the years 1990-1999 are used.

Keywords: Birth rates, Bayesian analysis, Count data, Generalized linear mixed models, Markov chain Monte Carlo, Teenage pregnancy.

1. Introduction

The teenage pregnancy rate in the USA is among the highest in the western nations. Four out of 10 American teens - nearly a million every year - become pregnant at least once before they turn 20. Teen pregnancy has serious social, health and financial consequences. As a social issue, the focus has been on reducing current teen pregnancy rates. Quite often the controversy is between the emotionally charged dichotomies of contraception and sex education versus abstinence.

Teen pregnancy rate declined fairly steadily from the late 1950's to mid 1980's and then climbed steeply between 1986 and 1991. Since then the rate has been showing a downward trend. Although the recent downward trend in teen pregnancy cuts across geographic, racial and ethnic lines, the drop has not been uniform in these subgroups. Teen pregnancy rate differential by race and ethnicity reflects disparities in education, income, access to medical care, concentration in poverty, family and community environment, and racial segregation. Differences at the individual level reflect different levels of sexual activity, use of contraceptives, availability of services and sex education, and attitudes about sexual behaviour. The factors associated with teen pregnancy operate at the individual as well as group (family, neighbourhood, community) level. The data at the individual level are not easily available but data at the community level (census districts, counties, states) are easily accessible. As a case study, we use teenage births counts and abortion counts data from the state of Texas.

Texas is one of the five most populous states with a population about 20 million with white, black and Hispanic residents in good proportions. This allows comparison among the three racial and ethnic groups. Texas 254 counties show a range of variation in population counts (from about one hundred to three million), urbanization levels, and social deprivation levels. Texan teenage birth rates are among the top 10 highest rates in the US.
The objectives of this report are

1. To study the socioeconomic and demographic factors influencing the variation in teenage birth rates among the white, black and Hispanic Texan girls,
2. To study the temporal and spatial trends in teen birth rates in Texas from 1990 to 1999.

2. Data

To model the birth rate, counties in Texas are used as the unit of aggregation. The data pertain to the 15-17 years old girls during the years 1991 to 1999. County level birth counts and population estimates data were gathered from the web pages of the Texas department of health (www.tdh.state.tx.us) and the Expert Health Data Programming Inc. of Houston, Texas (www.ehdp.com). County level data on socioeconomic, demographic and health services variables were gathered from various sources.

County level explanatory variables:

1. Urbancity: An index measuring the degree of urbanization in the county
2. Child Poverty: Percent of county residents under age 18 living under poverty in 1995
3. Not In School: Percent of 16-19 years olds not going to school and not high school graduates in 1990
4. FemLabor: Labor force female participation rate (percent) in 1990
5. FamClinics: The number of family planning clinics in the county in 1999

The birth rates per 1000 girls from the three race groups and over the years 1990-1999 are given in Figure 1. The racial differences in the rate decline are clearly evident. The decline of birth rates for black Texan teens is in line with the national trend.
3. Methodology

For analyzing birth counts, we consider a Poisson model. Let $Y_{it}$ denote the observed count of live births given by 15 to 17 years old girls in county $i$ ($1 \leq i \leq 254$) during the year $t$ ($1990 \leq t \leq 1999$). We assume that $Y_{it}$ is a Poisson random variable with mean $N_{it} \lambda_{it}$. Here $N_{it}$ and $\lambda_{it}$ are respectively the population size of 15-17 years old girls and the expected birth rate for this age group in county $i$ during year $t$. The Poisson distribution serves as an approximation to a binomial distribution, say $Y_{it} \sim \text{Bin}(N_{it}, \lambda_{it})$.

We relate the birth rate $\lambda_{it}$ to various covariates using a log-linear model

$$\log(\lambda_{it}) = X_i \beta + \gamma_{it} + b_i + \varepsilon_{it}$$

where,

- $\beta$: vector of parameters for county-specific covariates $X_i$,
- $\gamma_{it}$: spatio-temporal effects,
- $b_i$: county-specific random intercepts,
- $\varepsilon_{it}$: random errors.

Prior Distributions

Location parameters: $\beta_j \sim \text{independent N}(0,10000)$

Large-scale Spatio-temporal Variation

Quadratic temporal trend over each of the 11 public health regions (PHR’s)

$$\gamma_{0} = \gamma_{0|i} + \gamma_{1|i} t + \gamma_{2|i} t^2$$

$[ri] =$ Label of the PHR in which the county $i$ falls

$\gamma_{0|i}$ = PHR specific intercept
$\gamma_{1|i}$ = PHR specific linear trend coefficient
$\gamma_{2|i}$ = PHR specific quadratic coefficient

Prior Distributions:

- $\gamma_{0|i} \sim \text{N}(0, \sigma_{\gamma 0}^2)$; $\sigma_{\gamma 0}^2 \sim \text{IG}(0.0001,0.0001)$
- $\gamma_{1|i} \sim \text{N}(0, \sigma_{\gamma 1}^2)$; $\sigma_{\gamma 1}^2 \sim \text{IG}(0.0001,0.0001)$
- $\gamma_{2|i} \sim \text{N}(0, \sigma_{\gamma 2}^2)$; $\sigma_{\gamma 2}^2 \sim \text{IG}(0.0001,0.0001)$

Small-scale Spatial Dependence

Flexible CAR model: $b_i | b_{it} \sim N(\phi \bar{b}_i, \sigma_b^2)$; $i=1,...,254$

$\phi =$ spatial auto-regressive coefficient
$\bar{b}_i =$ mean value over the neighborhood of $i$
$\sigma_b =$ between-counties spatial dispersion.

Prior Distributions: $\phi \sim \text{Uniform}(-1,1)$; $\sigma_b^2 \sim \text{IG}(0.0001,0.0001)$

Residual Variation

AR(1) Model over time

$$\varepsilon_{i1990} \sim \text{N}(0, \sigma_{\varepsilon}^2), \varepsilon_{it} \sim N(\rho \varepsilon_{it-1}, \sigma_{\varepsilon}^2), t = 1991,...,1999$$

Prior Distributions: $\rho \sim \text{Beta}(2,2)$; $\sigma_{\varepsilon}^2 \sim \text{IG}(0.0001,0.0001)$
4. Results

Table 1 shows means and standard deviations of the marginal posterior distributions as obtained by MCMC computations using WinBugs software. The birth rates for all the groups go up with urbanization, with rising child poverty, and with the rise in the percentage of youngsters staying out of school. The impact of female labour participation on birth rates is not clear except for white teens.

Table 1. Posterior Marginal Summary of Regression Coefficients

<table>
<thead>
<tr>
<th>Covariate</th>
<th>All Races</th>
<th>Black</th>
<th>Hispanic</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Child Poverty</td>
<td>.0237</td>
<td>.0024</td>
<td>.0124</td>
<td>.0056</td>
</tr>
<tr>
<td>Urbanicity</td>
<td>.0093</td>
<td>.0012</td>
<td>.0081</td>
<td>.0030</td>
</tr>
<tr>
<td>Family Clinics</td>
<td>-.0158</td>
<td>.0034</td>
<td>-.0198</td>
<td>.0091</td>
</tr>
<tr>
<td>Not In School</td>
<td>.0092</td>
<td>.0026</td>
<td>.0107</td>
<td>.0064</td>
</tr>
<tr>
<td>Female Labor</td>
<td>.0030</td>
<td>.0034</td>
<td>.0016</td>
<td>.0075</td>
</tr>
</tbody>
</table>

The family planning clinics provide a variety of medical, educational and counseling services. It does appear that the availability of family planning services reduces the teen birth rate. To provide an interpretation of the regression coefficient value of –0.0158 for Family Clinics, it is estimated that each additional family planning clinic reduces the teen birth rate by about 1.6%. However, the effect of family planning clinics in reducing the birth rates is not uniform among the three races. Black and Hispanic teen birth rates are relatively less affected by the availability of the services. It could be due to the differential usage level by the racial groups or it could be due to non-availability of these services in predominantly black and Hispanic counties.

Table 2. Posterior Medians of Covariance Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All Races</th>
<th>Black</th>
<th>Hispanic</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>( \sigma^2_f )</td>
<td>0.052</td>
<td>0.057</td>
<td>0.067</td>
<td>0.079</td>
</tr>
<tr>
<td>( \sigma^2_b )</td>
<td>0.141</td>
<td>0.278</td>
<td>0.181</td>
<td>0.181</td>
</tr>
<tr>
<td>( \sigma^2_{\gamma_0} )</td>
<td>0.142</td>
<td>0.071</td>
<td>0.112</td>
<td>0.093</td>
</tr>
<tr>
<td>( \sigma^2_{\gamma_1} )</td>
<td>0.050</td>
<td>0.023</td>
<td>0.073</td>
<td>0.025</td>
</tr>
<tr>
<td>( \sigma^2_{\gamma_2} )</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td>\rho</td>
<td>0.922</td>
<td>0.883</td>
<td>0.847</td>
<td>0.876</td>
</tr>
<tr>
<td>\phi</td>
<td>0.824</td>
<td>0.459</td>
<td>0.911</td>
<td>0.935</td>
</tr>
</tbody>
</table>

Table 2 shows that posterior means for various covariance components are comparable over the three racial groups. There are, indeed, very strong spatial and temporal correlations. Birth counts for black females exhibit high inter-county variation and low spatial dependence. For all groups we observe that temporal trends show spatial variation, as measured by \( \sigma^2_{\gamma_0} \), \( \sigma^2_{\gamma_1} \) and \( \sigma^2_{\gamma_2} \).

Acknowledgements

The author thanks the Texas Department of Health and the Expert Health Data Programming Inc. of Houston for providing the data. This research is being supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.