Published in final edited form as:
Spat Spatiotemporal Epidemiol. 2017 June ; 21: 67–75. doi:10.1016/j.sste.2017.03.002.

Hot spots, cluster detection and spatial outlier analysis of teen birth rates in the U.S., 2003–2012

Diba Khan[a,*], **Lauren M. Rossen**[b], **Brady E. Hamilton**[b], **Yulei He**[a], **Rong Wei**[a], and **Erin Dienes**[c]

[a]Division of Research and Methodology/CDC/NCHS, Hyattsville, MD 20782, USA

[b]Division of Vital Statistics/CDC/NCHS, Hyattsville, MD 20782, USA

[c]RADARS SYSTEM/Rocky Mountain Poison & Drug Center (RMPDC)/A Division of Denver Health, 990 Bannock Street M/C 0180 | Denver, CO 80204, USA

Abstract

Teen birth rates have evidenced a significant decline in the United States over the past few decades. Most of the states in the US have mirrored this national decline, though some reports have illustrated substantial variation in the magnitude of these decreases across the U.S. Importantly, geographic variation at the county level has largely not been explored. We used National Vital Statistics Births data and Hierarchical Bayesian space-time interaction models to produce smoothed estimates of teen birth rates at the county level from 2003–2012. Results indicate that teen birth rates show evidence of clustering, where hot and cold spots occur, and identify spatial outliers. Findings from this analysis may help inform efforts targeting the prevention efforts by illustrating how geographic patterns of teen birth rates have changed over the past decade and where clusters of high or low teen birth rates are evident.

Keywords

Bayesian; Hotspots; Spatial outliers

1. Introduction

Teen childbearing is associated with negative health, social, and economic outcomes for both the mothers and infants (Ventura et al., 2014). Infants born to teen mothers are more likely to be born preterm, and to die within the first year of life compared to infants born to older mothers (Ventura et al., 2014). Moreover, the economic burden of teen childbearing to the public has been estimated to total 9.4 billion dollars annually in the U.S. (The National Campaign to Prevent Teen and Unplanned Pregnancy, 2013). At a national level, teen birth rates have evidenced a significant decline over the past few decades. In 2014, there were 24.2 births for every 1,000 adolescent females (15–19 years of age), a decline of nearly 60 percent over the past 25 years (Hamilton et al., 2014; Martin et al., 2013). Most of the states in the U.S. have mirrored this national decline, though some reports have illustrated

---

[*]Corresponding author at: Diba Khan, National Center for Health Statistics, Hyattsville, MD-20782. ild1@cdc.gov (D. Khan).

substantial variation in the magnitude of these decreases across the U.S. with declines ranging from 52% to 71% (Ventura et al., 2014). In addition, states in New England have historically had the lowest teen birth rates in the U.S., while several southern states (e.g., Arkansas, Mississippi, Oklahoma) have teen birth rates higher than the national average (The National Campaign to Prevent Teen and Unplanned Pregnancy, 2013; National Center for Health Statistics, 2014; Ventura et al., 2014).

Although many studies and reports have reported large-scale geographic patterns, spatial patterns at the county-level have not been much explored. Describing how counties with high or low teen birth rates cluster geographically may aid effort s to further reduce teen birth rates in specific areas of the U.S. Moreover, identifying counties that represent spatial outliers can inform future research seeking to better understand what factors might be driving greater success in reducing teen birth rates in some areas as compared to others. Many factors have been implicated in explaining the declines in teen birth rates such as increased contraceptive use and shifts toward more effective methods of contraception, delay in sexual initiation, the effects of the economic downturn, and the impact of various federal, state, and local public health programs and interventions to prevent teen pregnancy (Romero et al., 2015; Ventura et al., 2012). The objectives of this analysis were: to examine whether teen birth rates show evidence of clustering, to identify where hot and cold spots occur (groups of counties with extremely high or low teen birth rates), and to identify spatial outliers (counties with high or low teen birth rates surrounded by counties with dissimilar values), over the study period.

2. Methods

2.1. Data and estimates

Data on the number of live births for women aged 15–19 were extracted from the National Vital Statistics Birth Data Files for the years 2003–2012, (National Center for Health Statistics, 2003–2012). These data were then aggregated to the county level to provide a count of births to women 15–19 years of age for each county and year. The denominators to calculate teen birth rates were obtained from intercensal and postcensal population estimates of the number of females aged 15–19 years residing within each county over the same time period. These population denominators were extracted from the files containing intercensal and postcensal bridged-race population estimates provided by the National Vital Statistics System (National Vital Statistics System, 2009). There were 929 counties where there were fewer than 20 births to women 15 –19 years of age in 2003 and 1156 counties in 2012. Direct estimates are typically suppressed in counties with fewer than 20 cases in the numerator due to concerns about the stability and reliability of the estimates. To address these problems, we employed small area estimation methods to produce stable county-level estimates of teen birth rates from 2003 –2012. These methods are described elsewhere (Khan et al., 2018) and are reviewed briefly here. More detail can be found in the supplemental appendix.

Hierarchical Bayesian space-time interaction models were employed to produce county-level estimates of the teen birth rates for each year (Lawson, 2013). These models accounted for spatial and temporal dependence along with space time interaction terms to generate county-level estimates of teen birth rates for each year, 2003 –2012. These models included covariates related to county-level income and poverty, such as: per capita income, percent of county in poverty, unemployment rate, education, and racial/ethnic composition, proportion of foreign born residents, education level, measured at different time points. In addition to these covariates- which were significantly associated with teen birth rates at the county level, the number of family planning and Title X clinics by county, based on data provided by the Guttmacher Institute (Kost, 2014), were initially included in the models but subsequently removed due to the negligible association with teen birth rates at the county level. These models borrow strength across counties and time to produce stable estimates of teen birth rates at the county level, addressing problems due to data sparsity and allowing for further examination of geographic and temporal patterns.

2.2. Model

Let y_{it} = counts of teen births in county i and year t, and n_{it} = counts of teen population in county i and year t. Then,

$y_{it} \sim Binomial(n_{it}, p_{it})$; $i = 1,\ldots, m$ counties and $t = 1,\ldots, T$ years, where p_{it} = probability of teen births in county i at time t (Lawson, 2013; Khan et al., 2018).

The convolution model is:

$logit(p_{it}) = \alpha_0 + a_{1i} * year_t + X_i' \boldsymbol{\gamma} + u_i + \nu_i + \psi_{it}$. The components in the convolution model correspond to:

1. logit link function log ($p_{it}/(1 - p_{it})$).
2. α_0, an intercept.
3. time trend term $a_{1i}^* \; year_t$
4. $X_i' \boldsymbol{\gamma}$, where X_i: is the ith row of the covariates matrix and $\boldsymbol{\gamma}$ is a vector of regression parameters.
5. spatial random effects u_i by county to model strong spatial autocorrelation $i = 1, \ldots, m$ counties.
6. non-spatial random effects v_i by county to model residual spatial autocorrelation that were not dealt with by our spatial random effects, u_i, $i = 1,\ldots, m$ counties.
7. space-time interaction term ψ_{it}, a random effect where ψ_{it} is a function of its past values, $\psi_{i,t-1}$, plus an error term.

Parameters under (5) are modeled via normal conditional autoregressive priors (CAR) (Besag et.al, 1991). Parameters under (6) are modeled via normal conditional priors. Parameters under (7) are modeled via Type II random walk interaction (Knorr-Held and Rasser, 2000), which is included to account for any residual spatiotemporal dependence or variation that is not captured by the spatial or temporal main effects. The values for a given county in a given year depend upon the values observed for that county in the prior year plus a residual (Knorr-Held and Rasser, 2000; Lawson, 2013).

The predicted county-level teen birth rates obtained from the best model were merged with US Census Tiger/Line files and mapped using ArcGIS 10.1 (E.S.R. Institute, 2011).

3. Spatial statistical tools

Having obtained estimates from the models described above, several spatial statistical tools were implemented to examine spatial clustering and outliers.

3.1. Global index of spatial autocorrelation – Moran's I

Global indexes of spatial autocorrelation were used to assess the similarity, or spatial dependence, across counties with respect to teen birth rates. In other words, are counties with similar teen birth rates located close together or are teen birth rates randomly distributed across the US?

High values of the Moran's I and corresponding z-scores greater than 1.96 indicate that there is statistically significant clustering across the counties ($p < 0.05$). Low values of Moran's I and z-scores less than −1.96 indicate that there is statistically significant regularity (i.e., nearby counties have very different TBRs). Moran's I can be thought of as a spatially weighted form of Pearson's correlation coefficient (Waller and Gotway, 2004).

Global tests of autocorrelation such as Moran's I do not indicate where clusters of high or low teen birth rates might occur. Thus, local indicators of spatial autocorrelation (LISA) are needed.

3.2. Local indicators of spatial association – Getis-Ord Gi* and Anselin Local Moran's I

We used two LISAs to examine clusters of counties with high or low TBRs. First, the Getis-Ord Gi* statistic, which generates a z-score and corresponding p-value for each county, where z-scores greater than 1.96 indicate a significant "hot spot" and z-scores lower than −1.96 indicate a significant "cold spot" ($p < 0.05$).

Additionally, we used Anselin Local Moran's I to examine spatial outliers. This tool identifies spatial clusters of counties with high or low TBRs as well as spatial outliers. Spatial outliers refer to counties with values of TBRs that are discrepant from the neighboring counties. For example, a county with a particularly high TBR might be surrounded by counties with low TBRs. The local Moran's I index (I) is a relative measure and can only be interpreted within the context of its computed z-score or p-value (E.S.R. Institute, 2011). Based on these local Moran's I values and associated p-values, each county can then be classified by cluster/outlier type (COType). Counties fall into one of 5 groups (E.S.R. Institute, 2011):

- Not part of a cluster;
- Hot spot: High TBR and surrounded by a cluster of high TBRs (HH = "high-high");
- Cold spot: Low TBR and surrounded by a cluster of low TBRs (LL = "low-low");

- Spatial outlier: High TBR surrounded primarily by low TBRs (HL = "high-low");
- Spatial outlier: Low TBR is surrounded primarily by high TBRs (LH = "low-high").

3.3. Conceptualization of spatial relationships: defining a spatial weights matrix

There are several ways to define ω_{ij}, the binary spatial weight between county i and j. We explored two methods: K nearest neighbors and Delaunay triangulation. A general rule of thumb for K nearest neighbors is to evaluate each county in the context of a minimum of eight neighbors (E.S.R. Institute, 2011). If K (the number of neighbors) is 8, then the eight closest neighbors to the target county will be assigned positive nonzero weights, while counties outside of this boundary will receive a weight of zero. Sensitivity analyses were conducted using the Delanuay triangulation method of conceptualizing spatial relationships, where Voronoi triangles are drawn between county centroids and counties are defined as neighbors if they share a triangle edge. Delaunay triangulation is therefore a data driven approach to assigning the number of neighbors for a given county, in contrast to assigning an arbitrary number such as eight. Sensitivity analyses using this method of assigning spatial weights produced similar results to the 8-nearest-neighbors approach, thus we present only the latter. Distance based weighting schemes were not implemented because this approach leads to nearby counties being assigned very small weights, making it difficult to discern broader regional patterns.

4. Results

The predicted county-level teen birth rates for the years 2003 –2012 ranged from 6.1 –125.7 per 1000 in 2003 to 3.8–130.1 per 1000 in 2012. Figs. 1 and 2 depict the predicted teen birth rates across US counties for the years 2003 and 2012. Approximately 8.3 percent of counties had teen birth rates less than 20 per 1000 in 2003, while 14.6 percent of counties had teen birth rates less than 20 per 1000 in 2012.

4.1. Clustering of teen birth rates – Moran's I

The Global Moran's I and the corresponding z-score for the years 2003 –2012 suggest that there was significant spatial autocorrelation of county-level teen birth rates ($p < 0.05$) for all the years. In other words, across the US, counties with similar teen birth rates tend to locate closer to one another than we would expect by random chance during the period 2003 to 2012.

4.2. Hot and cold spots – Getis-Ord Gi*

Significant clusters of counties with high (hot spots) and low (cold spots) teen birth rates, as assessed by the Getis-Ord Gi* tool for the years 2003 and 2012 can be seen in Figs. 3 and 4. In 2003, 693 counties were part of a hot spot, and 760 counties were part of a cold spot. In 2012, the number of counties identified as within a hot spot declined to 633 and the number of counties that contributed to a cold spot decreased to 678.

4.3. Spatial outliers – Anselin Local Moran's I

Anselin Local Moran's I confirmed the significant hot and cold spots identified by the Getis Ord G * tool. In 2003, hot spots appeared across counties in the South such as southern Arizona and New Mexico, Arkansas, Florida, Georgia Louisiana, Mississippi, West Texas, and Appalachia. Cold spots were seen in Iowa, Michigan, Minnesota, Montana, North and South Dakota, Nebraska, New England, and Wisconsin. Generally, these patterns remained similar in 2012. The notable changes were in Arizona and counties across the South East (e.g., Alabama, Tennessee, Florida), which no longer represented a hot spot. Additionally, several counties were categorized as spatial outliers. These patterns can be seen for 2003 and 2012 in Figs. 5 and 6. Of note, in 2003 there were 30 counties identified as the HL ("high-low") type of spatial outlier, where the given counties evidenced high teen birth rates but were surrounded by a cold spot. Conversely, there were 14 counties identified as the LH ("low-high") type of spatial outlier, where teen birth rates were low, but the county was surrounded by a hot spot. In 2012, these numbers were relatively stable with 24 HL counties and 16 LH counties, respectively. In 2003, the HL counties were located in the states of Alabama, Arizona, Arkansas, California, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Maryland, Michigan, Missouri, Nebraska, Washington, and West Virginia. In 2012, HL counties appeared in these same states with the exception of four states that no longer contained HL counties: Illinois, Iowa, Kentucky and Virginia. In 2003, the LH counties were located in the states of Mississippi, North Carolina, North Dakota, Oregon, South Carolina, South Dakota, Tennessee, Texas, and Virginia. In 2012, the LH counties were located in these same states, with the exception of South Carolina and Virginia, which no longer contained LH counties. Additionally, two states, Pennsylvania and Wyoming, contained LH counties in 2012, but did not in 2003.

5. Discussion

There is substantial geographic variation in teen birth rates across the U.S. Results of global tests of spatial autocorrelation (i.e., Moran's I) confirm that teen birth rates exhibit spatial dependence. In other words, across the entire U.S., counties with high teen birth rates tend to locate closer together than we would expect at random. Conversely, counties with low teen birth rates also tend to cluster together geographically. Using local indicators of spatial association (i.e., Getis Ord Gi *), we were able to identify several hot and cold spots across the U.S. that represent clusters of counties with significantly high or low teen birth rates. The geographic pattern and change over time seen in the hot spots may reflect, in part, the wide differences in teen birth rates by race and Hispanic origin, as well as variations and changes in the composition of the teen populations by race and Hispanic origin across counties. A previous study of standardized teen birth rates by state in 2012 (Ventura et al., 2014) which controlled for compositional differences, nevertheless still showed regions of high and low rates. Even after accounting for racial/ethnic composition of the states, this prior report found that teen birth rates were highest in the southern states and lowest in the North Eastern states (Ventura et al., 2014), suggesting that there are other drivers of geographic patterns beyond racial/ethnic composition, such as area-level socioeconomic factors and other social determinants of health and health disparities. (Romero et al., 2016) The hierarchical Bayesian models employed in this analysis to predict TBRs included several county-level factors related to sociodemographic characteristics, socioeconomic factors, and racial/ethnic composition. (Khan et al., 2018) These factors were significantly related to TBRs at the county-level, adding to the body of literature describing the importance of social determinants of health in relation to teen childbearing and related disparities (Penman-Aguilar et al., 2013; Romero et al., 2016).

Consistent with prior research, we found higher teen birth rates across counties in the southern U.S. and lower teen birth rates in New England counties during the study period, 2003 to 2012. However, within those regions and states generally noted as having high or low teen birth rates, there is considerable variability at the county level, which crosses state borders in some cases. For example, clusters of high teen birth rates were seen for several counties along the border between west Texas and southeastern New Mexico, counties in Arkansas, Louisiana, Mississippi and Tennessee falling along the Mississippi river, some Appalachian counties in Kentucky and West Virginia, as well as counties along the border between southern Georgia and northern Florida. Similarly, cold spots emerged crossing state boundaries between Idaho and Montana, the Dakotas, Nebraska, Minnesota, Wisconsin and parts of Michigan, as well as nearly all of the New England states and into Pennsylvania, Maryland and northern Virginia. In this analysis, the presence and persistence of hot spots and cold spots are particularly noteworthy. Specifically, 695 counties were part of a hot spot in 2003, and 91.4% remained a hot spot in 2012. Conversely, 756 counties were part of a cold spot in 2003 and of these, 89.94% remained part of a cold spot in 2012.

This study has a few limitations. First, there may be variation in teen birth rates at the sub-county level, but this variation cannot be explored using data from the National Vital Statistics System. Second, the hierarchical Bayesian models used to generate the predicted annual teen birth rates at the county level were extremely computer intensive. The results of this analysis examining global and local indicators of spatial autocorrelation may be influenced by model error from the original hierarchical Bayesian model that was used to generate the smoothed county-level teen birth rates. Finally, the classification of counties as hot and cold spots or spatial outliers is based on their teen birth rate values relative to other counties in a given year. Thus, the stability of these specific categorizations does not reflect the magnitude of declines in teen birth rates across the U.S., but only the relative stability in the ‘ranking’ of counties from high to low. The strength of this analysis is the combination of a detailed geographic focus (at the county-level), over a substantive period of time. To date, most of the work on estimating teen birth rates and assessing geographic variation has been done at the state level (Ventura et al., 2014). Examinations of teen birth rates at the county level have relied on direct estimates aggregated over several years, and are limited by the inability to show county-level estimates based on fewer than 20 births in the numerator due to concerns about the reliability and stability of the direct estimates. (Romero et al., 2016) This study is the first study to describe county-level variation in teen birth rates over the entire U.S., by using estimates derived from Hierarchical Bayesian models and applying spatial statistical tools to identify hot spots, cold spots, and spatial outliers in teen birth rates.

6. Conclusions

In sum, there is substantial geographic variation in teen birth rates across the U.S. Counties with high and low teen birth rates tend to cluster together more than we would expect by chance. Describing how counties with high or low teen birth rates cluster geographically may aid effort s to further reduce teen birth rates in specific areas of the U.S. Findings may be used by public health practitioners and other officials to inform decisions related to the development and implementation of teen pregnancy prevention effort s by facilitating better targeting of resources toward areas in greatest need. Additionally, results related to spatial outliers may inform future research aimed at understanding what may be contributing to lower or higher than expected TBRs (i.e., based on dissimilar values in adjoining counties) in specific counties identified as a HL or LH spatial outlier. Despite significant declines in teen birth rates at the national level, there are many counties and groups of counties across the US where teen birth rates remain high. Identifying counties that represent spatial outliers can inform future research seeking to better understand what factors might be driving greater success in reducing teen birth rates in some areas as compared to others. Given differences in teen birth rates across racial/ethnic subpopulations and specific age ranges such as 15–17 and 18–19, future research might look at county-level birth rates for women aged 15–17 and 18–19 and for race groups such as non-Hispanic white, non-Hispanic black, and Hispanic women aged 15–19 to examine hot spots and spatial outliers which can provide additional insight about geographic patterns and potential drivers of these patterns and trends.

References

Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math. 1991; 43:1–20.

E.S.R. Institute. ArcGIS desktop: release 10. Redlands CA: Environmental Systems Research Institute; 2011. Available from: http://resources.arcgis.com/en/help/main/10.1/index.html#/Modeling_spatial_relationships/005p00000005000000/

Hamilton, BE., Martin, JA., Osterman, MJ., Curtin, SC. Preliminary data for 2014. Hyattsville, MD: National Center for Health Statistics; 2014. National Vital Statistics Report; p. 64p. 6

Khan, D., Rossen, LM., Brady, H., Erin, D., He, Y., Wei, R. Spatiotemporaltrends in teen birth rates in the U.S., 2003–2012. J R Stat Soc Ser A. 2018. http://onlinelibrary.wiley.com/doi/10.1111/rssa.12266/full

Knorr-Held L, Rasser G. Bayesian detection of clusters and discontinuities in disease maps. Biometrics. 2000; 56:13–21. [PubMed: 10783772]

Kost, K. Guttmacher county-level contraceptive needs and services data. New York: The Alan Guttmacher Institute (AGI); 2014.

Lawson, A. Bayesian disease mapping hierarchical modeling in spatial epidemiology. Chapman and Hall, CRC; 2013.

Martin JA, Hamilton BE, Ventura SJ, Osterman MJ, Mathews T. Births: final data for 2011. Nat Vital Stat Rep. 2013; 62:1–90.

National Center for Health Statistics. Nation at a glance: declining teen birth rates by state. 2014. Available from http://www.cdc.gov/nchs/features/nation_oct2014/nation_at_a_glance_oct2014.htm

National Center for Health Statistics; 2003–12. Natatlity data, 2003–2012, Geographic detail (all counties) file

National Vital Statistics System. Bridged-raceintercensal population estimates. 2009. Available from https://www.cdc.gov/nchs/nvss/bridged_race/data_documentation.htmUse

Penman-Aguilar A, Carter M, Snead MC, Kourtis AP. Socioeconomic disadvantage as a social determinant of teen childbearing in the u.s. Public Health Rep. 2013; 128(Suppl 1):5–22. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5306a6.htm.

Romero, L., Pazol, K., Warner, L., et al. Morb Mortal Wkly Rep. 2015. p. 64Available from http://www.cdc.gov/mmwr/pdf/wk/mm64e0407.pdf

Romero L, Pazol K, Warner L, et al. Reduced disparities in birth rates among teens aged 15–19 years united states, 2006–2007 and 2013–2014. Morb Mortal Wkly Rep. 2016; 65(16):409–14.

The National Campaign to Prevent Teen and Unplanned Pregnancy Teen birth rate comparison. 50 state comparison story. 2013; 2013 Available from: https://thenationalcampaign.org/data/compare/1701.

The National Campaign to Prevent Teen and Unplanned Pregnancy. Counting it up: The public costs of teen childbearing: Key data. Available from https://thenationalcampaign.org/resource/counting-it-key-data-2013

Ventura, SJ., Curtin, SC., Abma, JC., Henshaw, SK. National Vital Statistics Reports 2012. Hyattsville, MD: National Center for Health Statistics; 2012. Estimated pregnancy rates and rates of pregnancy outcomes for the united states, 1990–2008.

Ventura, SJ., Hamilton, BE., Mathews, TJ. National Vital Statistics Reports 2014. Hyattsville, MD: National Center for Health Statistics; 2014. National and state patterns of teen births in the united states, 1940–2013.

Waller, L., Gotway, C. Applied spatial statistics for public health data. New York: John Wiley and Sons; 2004.

Appendix

This section is adapted from Khan et al. (2018).

The convolution model is:

$$logit\,(p_{it}) = \alpha_0 + a_{1i} * year_t + X_i{}' \boldsymbol{\gamma} + u_i + v_i + \psi_{it},$$

where p_{it} is the probability of teen birth for county i at time t is modeled using a logit link. The county level covariates accounting for the term $X_i{}' \boldsymbol{\gamma}$, in the convolution model were obtained from Area Resource Files (The National Campaign to Prevent Teen and Unplanned Pregnancy, 2013). The covariates included various socioeconomic indicators (e.g., per capita income, percent of county in poverty, unemployment rate) and demographic variables (e.g., racial composition, proportion of foreign born residents, education level), measured at different time points. To account for collinearity within the predictors and to reduce and transform data in the presence of multicollinearity, a Principal Component Analysis (PCA) was conducted. The PCA, using a varimax rotation, indicated that three three components were sufficient as they accounted for 88% of the total variation. The three components largely reflected county level income and poverty, education, and percent non-Hispanic white, respectively. More details are available from National Center for Health Statistics (2014).

Prior assumptions

The overall intercept term, α_0 is assigned an improper flat prior. The prior specifications for all other components in the model are as follows:

- $\alpha_0 \sim dflat\,()$ (A.1)

- Prior for $a_1 \sim N(0, \tau_{a_1})$ (A.1)

  where, $\tau_{a_1} \sim IG(0.01, 0.01)$, (IG:Inverse Gamma)(A.2)

- $\psi_{i,t} \sim N(\psi_{i,t-1}, \tau\psi)$ (A.3),

  where $\psi_{i,1} \sim N(0, \tau_{\psi})$ (A.4), and

  $\tau_{\psi} \sim IG(0.01, 0.01)$(A.5).

  The prior for Type II random walk interaction is defined above and can be regarded as a form of residual [(Knorr-Held and Rasser, 2000; Lawson, 2013)].

- $\gamma \sim N(0, \tau_{\gamma})$ (A.6), where

  $\tau_{\gamma} \sim IG(0.01, 0.01)$ (A.7)

  Intrinsic Conditionally Autoregressive Prior (CAR) prior for $u_i|u_{-i}$:

- $u_i|u_{-i} \sim N(\overline{u_i}, \tau_u/n_{\delta_i})$ (A.8) termed as correlated heterogeneity (variability), where,

  $u_{-i} = (u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m)$

  $\overline{u_i} = \frac{\sum_{j=1}^{m} \omega_{ij} u_j}{\sum_{j=1}^{m} \omega_{ij}}$ (A.9)

  δ_i: neighborhood of i^{th} region

  n_{δ_i}:number of neighbors$=\sum_{j=1}^{m} \omega_{ij}$

  $\omega_{ij} = 1$ for counties i and j that are deemed neighbors otherwise 0.

  We used the K-nearest neighbors approach to define the spatial relationships between counties. The eight closest neighbors to the target county were assigned weights $\omega_{ij} = 1$, all other counties are assigned weights $\omega_{ij} = 0$.

  τ_u : is the variance. $\tau_u \sim IG(0.01, 0.01)$ (A.10)

- Prior for v_i: $v_i \sim N(0, \tau_v)$ (A.11)

  termed as uncorrelated heterogeneity (variability), where

  τ_v : is the variance

  τ_v is assigned the prior $\sim IG(0.01, 0.01)$ (A.12)

  Prior Assumptions: Non-Spatial Model $\boldsymbol{\beta}_i = (\beta_{1i}, \beta_{2i})'$ (A13)

  $\beta_i \overset{iid}{\sim} N(\boldsymbol{\mu}, \boldsymbol{\Upsilon})$ (A14)

  Hyper prior for $\boldsymbol{\mu}$ is assumed as $\sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{pmatrix}\right)$ (A15)

  The inverse of $\boldsymbol{Y}$ is assigned a Wishart prior with $\boldsymbol{R}$ as a diagonal matrix and degrees of freedom = 2.

Inverse of $$(\boldsymbol{\Upsilon}) \sim Wish\left(\begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}, 2\right) \quad \text{(A16)}$$

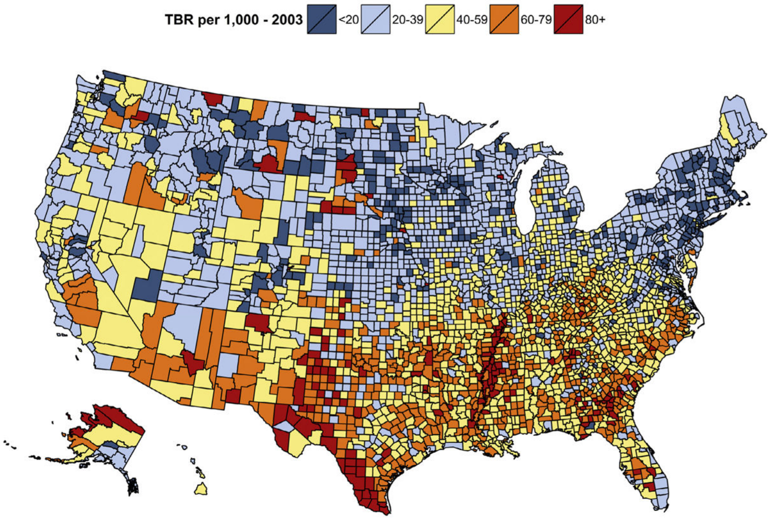

Fig. 1.

Model-based estimates of teen birth rates (per thousand) for 2003.

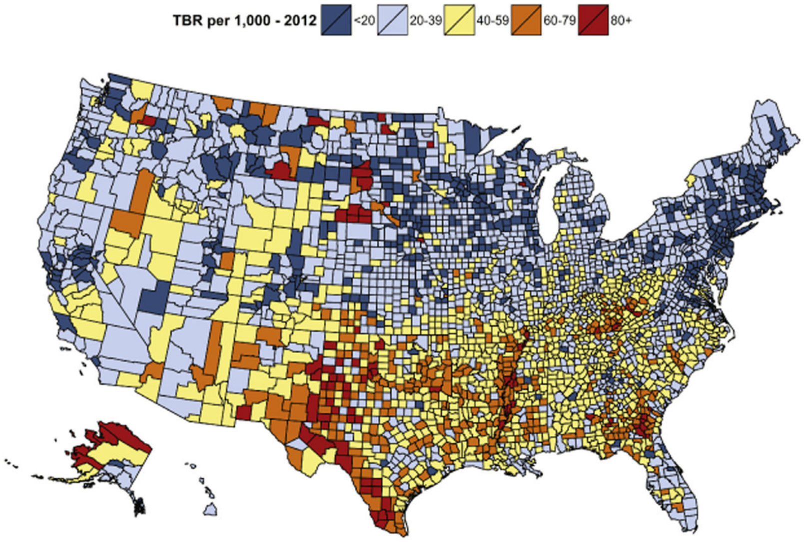

Fig. 2.

Model-based estimates of teen birth rates (per thousand) for 2012.

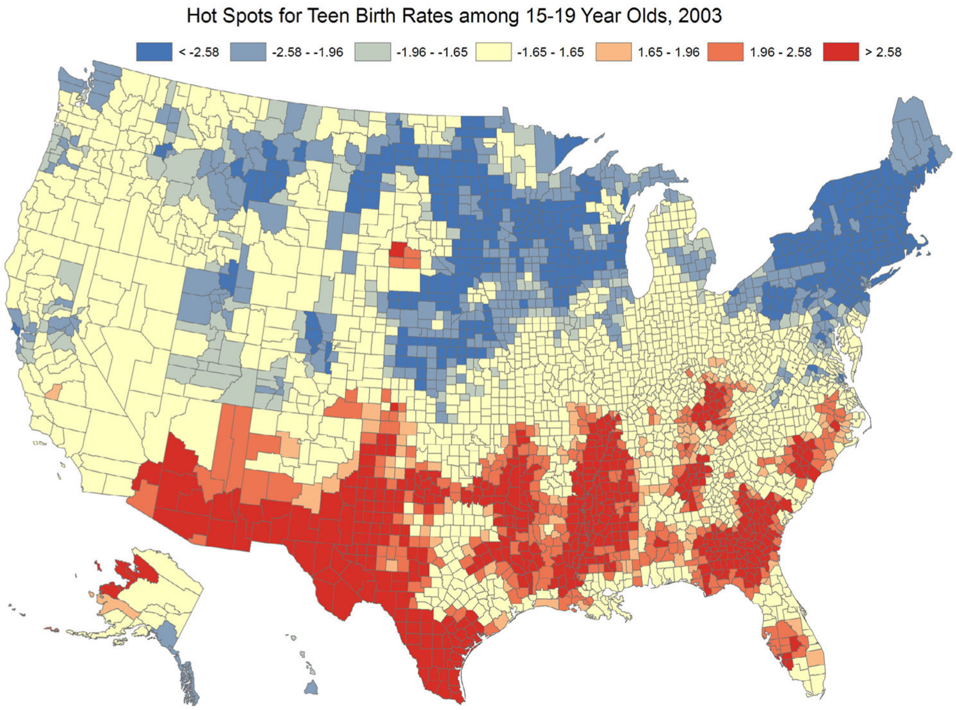

Fig. 3.

Hot and Cold Spots in Teen Birth Rates, 2003. Values represent z-scores from the Getis Ord Gi * analysis; 1.65 corresponds to $P < 0.10$, 1.96 corresponds to $P < 0.05$, 2.58 corresponds to $P < 0.01$. Negative z-scores indicate cold spots, while positive z-scores indicate hot spots.

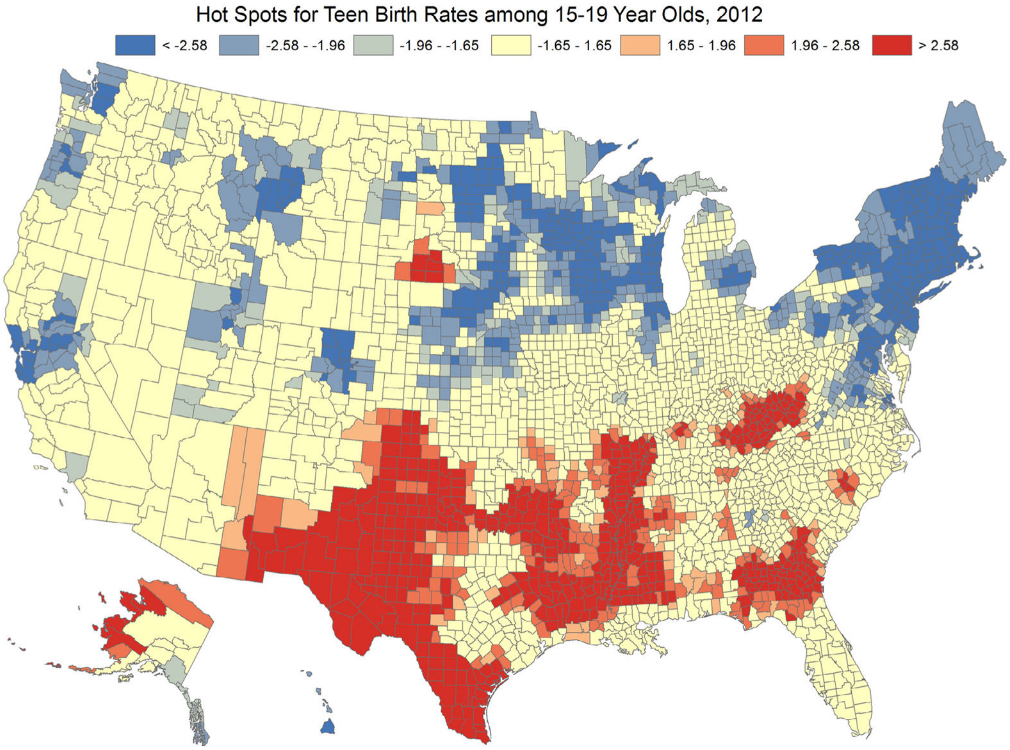

Fig. 4.

Hot and Cold Spots in Teen Birth Rates, 2012. Values represent z-scores from the Getis Ord Gi* analysis; 1.65 corresponds to $P < 0.10$, 1.96 corresponds to $P < 0.05$, 2.58 corresponds to $P < 0.01$. Negative z-scores indicate cold spots, while positive z-scores indicate hot spots.

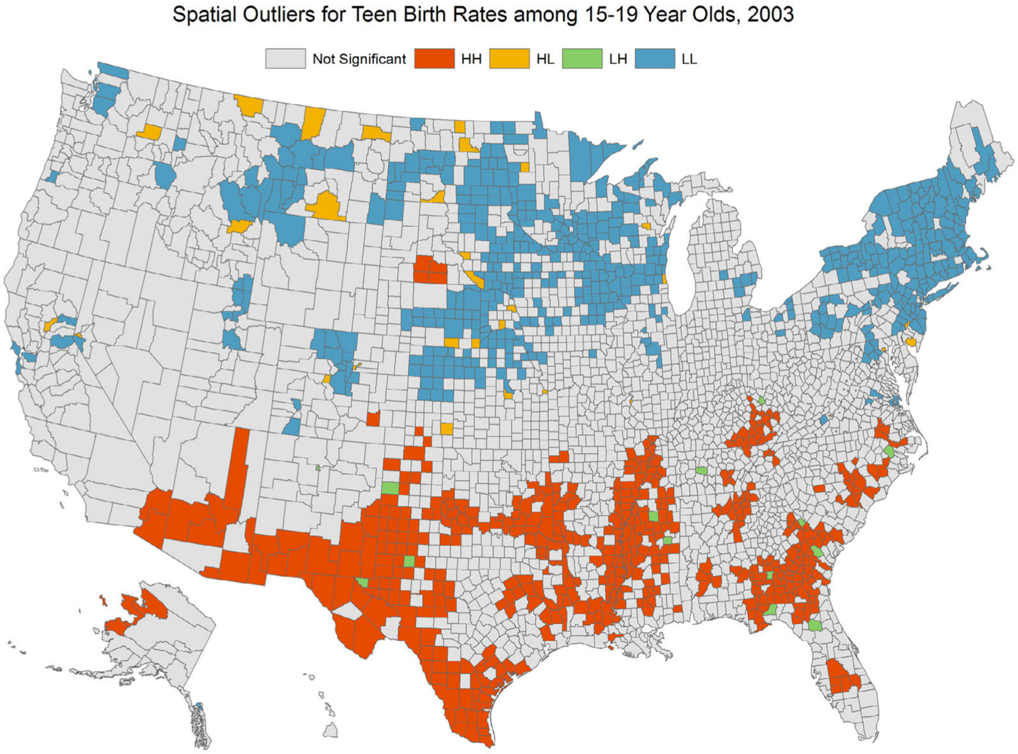

Fig. 5.

Clusters and spatial outliers in teen birth rates, 2003.

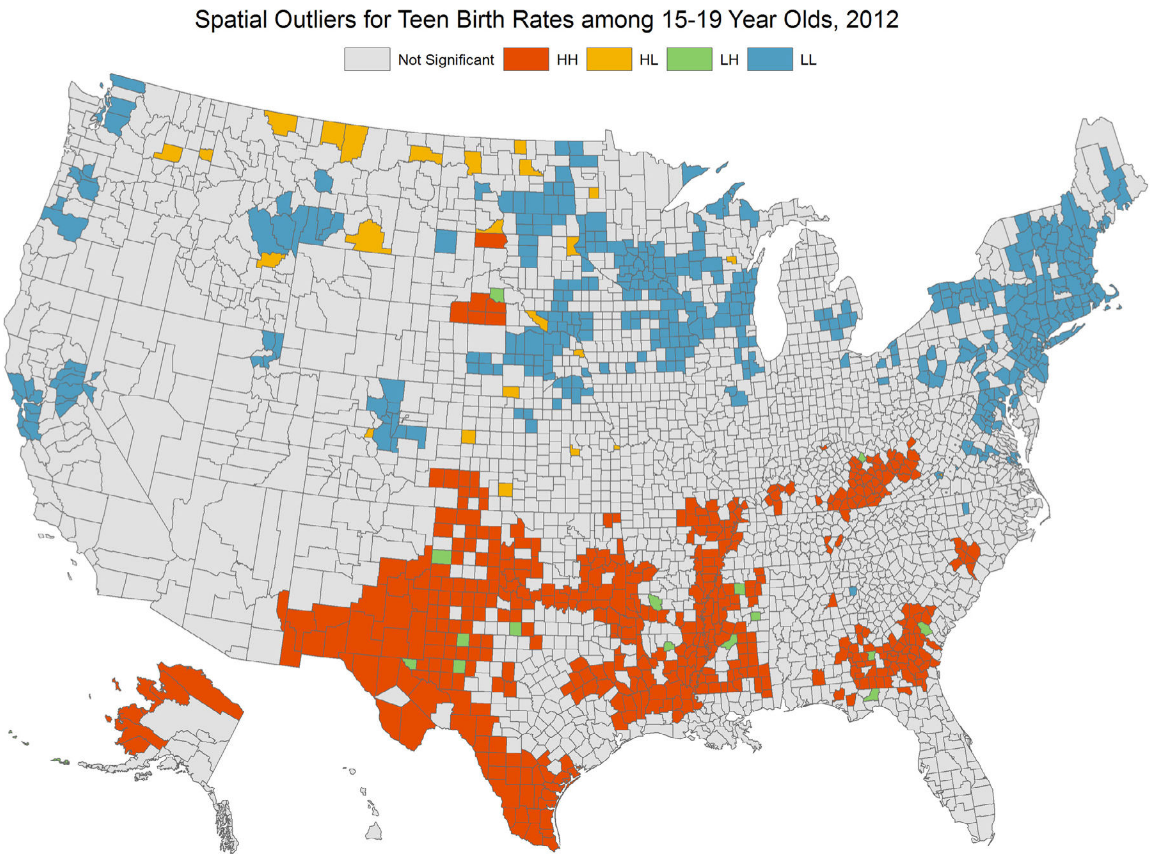

Fig. 6.

Clusters and spatial outliers in teen birth rates, 2012.