Geographic and community factors associated with adolescent pregnancy in New York State

Elizabeth Villamil
University at Albany, State University of New York, exv02@health.state.ny.us

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Epidemiology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/794

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
GEOGRAPHIC AND COMMUNITY FACTORS ASSOCIATED WITH ADOLESCENT PREGNANCY IN NEW YORK STATE

by

Elizabeth Villamil

A Dissertation Submitted to the University at Albany, State University of New York in Partial Fulfillment of the Requirements for the Degree of Doctor of Public Health

School of Public Health
Department of Health Policy, Management and Behavior

2012
GEOGRAPHIC AND COMMUNITY- FACTORS
ASSOCIATED WITH
ADOLESCENT PREGNANCY IN NEW YORK STATE

by

Elizabeth Villamil

COPYRIGHT 2012
ABSTRACT

GEOGRAPHIC AND COMMUNITY-FACTORS ASSOCIATED WITH

adolescent pregnancy in new york state

Background

Examining geographic variation in adolescent pregnancy can help identify important community factors that are linked by geography and generate hypotheses about the underlying causes of community disparities. In this study, a spatial scan statistic method, combined with a generalized linear mixed model were used to determine whether adolescent pregnancy varies by place of residence, after adjusting for community factors, and whether certain communities can be targeted as at higher risk for adolescent pregnancy in the future.

Methods

Using population-based and vital records data, adolescents between 15-19 years, residing in New York State who gave birth from 2006-2008 were analyzed. Clusters of communities with high adolescent pregnancy prevalence were adjusted for community factors (race, ethnicity, education, poverty and single parent households) and evaluated using a spatial scan statistic to identify places where adolescent pregnancy prevalence was significantly higher than the statewide experience. To determine risk factors associated with clusters of communities with a high prevalence of adolescent pregnancy a generalized linear mixed model was employed.
Results

Results indicate that nine communities experienced statistically significant excess adolescent pregnancy prevalence. After adjusting for community factors education, race, ethnicity, poverty and single parent households, individually, geographical clustering of communities with high adolescent pregnancy prevalence persisted. Four clusters emerged after adjusting for the community factor single parent households. When adjusting for the six community factors together, all significant clustering of communities with high adolescent pregnancy prevalence disappeared. The generalized linear mixed model showed that, poverty, single parent households, ethnicity, foreign born citizens and race, were all positively associated with adolescent pregnancy.

Conclusion

Race, education, ethnicity, poverty and single parent households, individually, are not a main contributing community factor for risk of adolescent pregnancy. Conversely, race, ethnicity, education, poverty and single parent households, combined, may be contributing causes of community disparities. Clusters of communities in New York State are experiencing high adolescent pregnancy prevalence; however, there are many more communities at risk for high adolescent pregnancy prevalence. The problem of adolescent pregnancy also has the expectation of moving from a primarily urban city issue to a concern of poorer rural areas.
DEDICATION

I dedicate this dissertation to Nanny and Poppy, for their unconditional love and support.
ACKNOWLEDGEMENTS

I wish to express my most sincere gratitude to everyone who has supported me in this endeavor. In particular, I would like to thank my committee chair, Dr. Benjamin Shaw, who encouraged and challenged me throughout this process. His diligence enabled me to develop this dissertation from an initial idea all the way to completion. I would like to thank Dr. Francis Boscoe for his precious time and generous guidance on the statistical aspect of this dissertation. I am grateful for the good advice, support and friendship of my committee member Dr. Barry Sherman. Dr. Sherman has been invaluable on both an academic and a personal level. I would like to thank the many other people who have contributed to my education and research experiences. I would like to thank my colleagues at the School of Public Health for their continuous support and encouragement. Lastly, I would like to thank my family and friends for all of their love and support.
**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td>1.1. Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Significance</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Purpose of the Study</td>
<td>3</td>
</tr>
<tr>
<td>1.4. Conceptual Framework</td>
<td>5</td>
</tr>
<tr>
<td>1.5. Research Questions and Hypothesis</td>
<td>8</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td></td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.2. Adolescence and Sexuality</td>
<td>10</td>
</tr>
<tr>
<td>2.3. Pregnancy among Adolescents</td>
<td>11</td>
</tr>
<tr>
<td>2.4. Antecedents of Adolescent Pregnancy</td>
<td>12</td>
</tr>
<tr>
<td>III. RESEARCH DESIGN AND METHODOLOGY</td>
<td></td>
</tr>
<tr>
<td>3.1. Geographic Unit of Analysis</td>
<td>23</td>
</tr>
<tr>
<td>3.2. Case Finding and Data Collection Methods</td>
<td>25</td>
</tr>
<tr>
<td>3.3. Variables</td>
<td>31</td>
</tr>
</tbody>
</table>
### IV. DATA ANALYSIS

4.1. Spatial Analysis – Aim One and Aim Two 33  
4.2. Community Analysis Modeling Framework-Aim Three 37  
4.3. Geographical Mapping 40

### V. RESULTS

5.1. Descriptive Analysis 41  
5.2. Spatial Analysis Results – Aim One and Two 43  
5.3. Community Analysis – Aim Three 63

### VI. DISCUSSION AND CONCLUSION

6.1 Aim One 69  
6.2 Aim Two 73  
6.3 Aim Three 80  
6.4 Future Recommendations 84  
6.5 Limitations 85  
6.6 Conclusion 86

### REFERENCES 89
## LIST OF TABLES

### Chapter 5

| Table 5.1 | Statistically significant clusters of adolescent pregnancy at the ZIP code level in New York State, 2006-2008 |
| Table 5.2 | Tracking results for part of the City of Albany to confirm factors of most importance, after adjustments |
| Table 5.3 | Tracking results of part of the City of Niagara Falls to confirm factors of most importance, after adjustments |
| Table 5.4 | Tracking results of part of the City of Binghamton to confirm factors of most importance, after adjustments |
| Table 5.5 | Tracking results of part of the City of Watertown to confirm factors of most importance, after adjustments |
| Table 5.6 | Beta coefficients and p values resulting from the generalized linear mixed model |
LIST OF FIGURES

Chapter 1

Figure 1.1 Bronfenbrenner’s Ecological Model

Chapter 5

Figure 5.1. Adolescent pregnancy rates, aggregated from 2006-2008

Figure 5.2. Statistically significant clusters of communities with high adolescent pregnancy prevalence.

Figure 5.3. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for education

Figure 5.4. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for race

Figure 5.5. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for ethnicity

Figure 5.6. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for poverty

Figure 5.7. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for single parent households

Figure 5.8. Comparison of adolescent pregnancy by ZIP code from 2000 – 2010

Figure 5.9. New York State communities observing an increase in the level of risk for adolescent pregnancy, from 2000 to 2010.
CHAPTER I

INTRODUCTION

1.1 Background

The prevention of unintended adolescent pregnancy is an important goal of our society. For several decades, concerns have been mounting in the United States regarding adolescent sexual behavior, pregnancy and parenthood. Sexually active teenagers are at increased risk for unintended pregnancies. Adolescents, because they are in developmental transition, are particularly sensitive to environmental and contextual influences. Health outcomes for adolescents are grounded in their social environments and are frequently influenced at the community level. Therefore, if environmental factors challenge adolescents’ decision-making and health, the burden of unwanted pregnancies rest not only on the adolescent, but also on society as a whole.

Adolescent pregnancy rates in the United States rose steadily from the 1970s to the 1990s, increasing by approximately 21 percent among all females younger than 20 during those two decades. From 1990 to 2004, widespread declines in births among adolescents were reported nationwide. In 2005, the average birth rates for 15-19 year olds were 40.5 births per 1,000 women. In 2006, average birth rates increased to 41.9 births per 1,000 women. Approximately 10 percent of all U.S. births in 2006 were to adolescents. Roughly 51 percent of adolescent pregnancies end in live births, 35 percent end in induced abortion, and 14 percent result in miscarriage or stillbirth. Seventy-eight percent of these pregnancies are unplanned, accounting for roughly one-quarter of all accidental pregnancies annually. This is cause for concern, as teenage
pregnancy can have substantial economic, societal, and health repercussions. The 2009 New York State Risk Behavior Survey indicated an average of 42 percent of all high school students grades 9 through 12 (9th grade 26.4 percent, 10th grade, 37.3 percent, 11th grade, 46.2 percent and 12th grade, 61.8 percent) had sexual intercourse.\(^7\)

The total annual government expenditures on public aid to adolescent mothers in 2006 were $11.3 billion.\(^8\) Adolescent childbearing in New York State cost taxpayers roughly $377 million in 2008, according to an analysis from The National Campaign to Prevent Teen and Unplanned Pregnancy.\(^9\) Of these costs, 24 percent were federal costs and 76 percent were state and local costs. For the nation overall, adolescent childbearing costs taxpayers $10.9 billion.\(^9\) The majority of costs of adolescent childbearing are associated with negative consequences for the children of teen mothers, including increased costs for health care, foster care and incarceration. Annual taxpayer costs associated with children born to teen mothers include public health care (Medicaid and Child Health Plus), child welfare, and, among those children who have reached adolescence and young adulthood, increased rates of incarceration, and lost tax revenue due to decreased earnings and spending.\(^9\)

1.2 Significance

The statistics detailed above underscore the necessity for comprehensive prevention efforts that include programming related to adolescent pregnancy prevention. There is no shortage of opinions as to what will reduce adolescent pregnancy, nor is there a shortage of program models or model programs. Researchers have examined the relationship between individual characteristics and measures of sexual activity,
contraceptive use, pregnancy and childbearing.\textsuperscript{10,11} Findings demonstrate a wide range of antecedents related to teenage sexual risk-taking and childbearing. However, merely examining individual level factors related to adolescent pregnancy limits understanding of these complex, multi-dimensional public health issues, and prevents insight into how multiple level factors influence sexual behavior and health.\textsuperscript{12} Increasingly, researchers and epidemiologists are recognizing important roles that community contexts and structural factors play in determining adolescent pregnancy and other health risks.\textsuperscript{12} In particular, it is important to consider the social context in which individual factors operate. The characteristics of the community setting in which an adolescent resides provide a context that has the potential to influence adolescent pregnancy and to inform the development and expansion of accessible and effective community-based prevention and treatment services.

1.3 	extbf{Purpose of the Study}

The purpose of this study is to identify statistically significant clusters of communities (represented by ZIP codes) with a high prevalence of adolescent pregnancy that either persist or disappear after adjustment for race, ethnicity, education, single parent household or poverty. Additionally, this study seeks to determine the influence and interaction of community factors on adolescent pregnancy. While all adolescents are at risk, some adolescents are at increased risk for early sexual activity, poor contraceptive use and pregnancy. Numerous risk factors (community, peer, family and culture) have been shown to influence adolescent sexual behavior. One of the missions of public health is to identify and define models to reduce these risks, to prevent adolescent
pregnancy. Knowing whether adolescent pregnancy varies geographically is especially relevant since area-based physical, social and behavioral factors can assist or impede the effectiveness and efficiency of these prevention models. The objective of this research is to provide timely and relevant information to public health professionals to assist in identifying New York State communities currently observing a high prevalence, and to identify communities which may expect to see a higher prevalence of adolescent pregnancy, based on underlying community factors. Armed with this information, public health professionals will be able to develop efficient, effective and targeted interventions to reduce the number of pregnancies among adolescents in New York State.

The specific aims of this study are as follows:

**Aim One**: Identify any statistically significant clusters of communities (represented by ZIP codes) with a high prevalence of adolescent pregnancy.

**Aim Two**: To determine if community factors (race, ethnicity, single parent households, education and poverty) are associated with communities of high adolescent pregnancy prevalence.

**Aim Three**: To determine whether the risk for adolescent pregnancy prevalence changed among communities from 2000 to 2010.
1.4 Conceptual Framework

The purpose of this research is to use Urie Bronfenbrenner’s ecological systems model to identify factors to successfully predict which New York State communities have a high prevalence of adolescent pregnancy. The influence of various community factors on the developing adolescent impacts whether an adolescent engages in risky sexual behaviors, potentially leading to adolescent parenthood. Bronfenbrenner’s model recognizes the interwoven relationships between people and their living environments. While individuals are responsible for instituting and maintaining the lifestyles necessary for their health and personal behaviors, their health consequences are determined to a large extent by their social and physical environment, such as living conditions, community norms and values, environment, health services, regulations, and policies. This ecological model, because of its inherent focus on context as well as the individual, has gained considerable prominence as an approach to study adolescent risky sexual behavior and adolescent pregnancy.13

Bronfenbrenner’s seminal work, The Ecology of Human Development (1979), identified four system levels: the microsystem; mesosystem; exosystem; and, the macrosystem, which interact multi-directionally (Figure 1.1).14 At the most basic interactional level is the microsystem, these factors operate at the individual level, and include an adolescent’s attitudes, beliefs, future expectations, knowledge and behaviors. The relationships inside the microsystem have bi-directional influences, in that they impact in two directions, both away and toward the developing adolescent. The mesosystem, includes the adolescent’s family structure, parent-child communication and peer influences. The exosystem, of most relevance to this study, includes characteristics
of an adolescent’s community, including school context, neighborhood environment and the norms and values concerning adolescent pregnancy. The last level, the macrosystem, includes the society, culture and broad policies in which an adolescent develops.

Bronfenbrenner’s model can be used to assess the complex and multi-determined nature of the social problem of adolescent pregnancy with its conceptualization of the individual as being affected by the larger system levels. A central tenet in Bronfenbrenner’s framework implies that the greatest influence on development is from the outside to the inside; meaning, influence is most salient from the major culture through the exosystem and mesosystem, to the microsystem, otherwise known as the developing adolescent. Bronfenbrenner argues that social interactions, like biological interactions, are always part of a larger ecological system. Bronfenbrenner advocates for the necessity to expand a research perspective to examine multi-person systems interaction, taking into account aspects of the community beyond the immediate situation containing the individual.

Following this ecological model, the research conducted for this dissertation identifies the prevalence of adolescent pregnancy within the state and examines the area-based social determinants. This study uses two main levels of statistical analyses: spatial and area-based. The spatial analysis is used to identify statistically significant clusters of communities with a high prevalence of adolescent pregnancy, designating these distinctions by geographic location on a map. Area-based analysis is used to explore the potential interaction of community factors and determine how these factors geographically align with the statistically significant clusters of adolescent pregnancy.
Figure 1.1. Bronfenbrenner’s Ecological Model

MACROSYSTEM
society, political systems,
culture, nationality, economics

EXOSYSTEM
community, neighborhoods,
school

MESOSYSTEM
family, friends,
social network

MICROSYSTEM
behaviors, attitudes,
knowledge, skills
1.5 Research Questions and Hypotheses

The following research questions guide the study:

- Do statistically significant clusters of communities (represented by ZIP codes) with a high prevalence of adolescent pregnancy exist in New York State?
- Do clusters of communities with a high prevalence of adolescent pregnancy in New York State persist or disappear after adjusting for race, ethnicity, single parent households, education or poverty?
- Has the risk for adolescent pregnancy prevalence changed among communities from 2000 to 2010?

The research questions above will be answered using the following null hypotheses:

**Hypothesis 1:** No statistically significant spatial clusters of adolescent pregnancy exist in New York State communities.

**Hypothesis 2:** Clusters of communities with a high prevalence of adolescent pregnancy persist after adjusting for race, ethnicity, single parent households, education and poverty.

**Hypothesis 3:** There will be no difference in risk for adolescent pregnancy prevalence among communities from the year 2000 to 2010.
CHAPTER II

REVIEW OF RELATED LITERATURE

2.1 Introduction

Literature on adolescent sexual behavior is examined from a multisystem perspective. Such an approach is guided by Bronfenbrenner’s Ecological Systems Model, which emphasizes the reciprocal relations among multiple systems of influence on a person’s behavior. According to this perspective, an accurate and comprehensive understanding of adolescent sexual risk behavior must include some knowledge of both the personal and environmental factors which may contribute to the decision to become sexually active, and subsequently, the decision to engage in risk-promoting or risk-reducing sexual behavior.

To gather relevant articles for this review, several search methods were utilized. First, two on-line databases in the social and health sciences (PubMED and ERIC) were searched for publications dated between 1990 and 2010, using the following terms: adolescent sexual behavior, adolescent sexual activity, adolescent sexuality, adolescent risky sexual behavior and adolescent pregnancy. The initial search produced more than 600 articles, which were significantly reduced after applying the following criteria: 1) participants in reported study must be less than 19 years of age; 2) publications were from peer-reviewed journals; and, 3) study included measures of adolescent sexual risk behavior or adolescent pregnancy as an outcome. Adolescent sexual risk behavior was defined as inconsistent or non-use of condoms, inconsistent or non-use of other contraceptive methods, having multiple sexual partners, and substance abuse in conjunction with participation in sexual activity. These behaviors were chosen as
indexes of sexual risk behaviors because they are all well represented among the outcomes in the adolescent sexual behavior literature and clearly place adolescents at risk for negative outcomes such as pregnancy or sexually transmitted diseases. Articles from bibliographic reviews of the acquired articles and manual searches of relevant journals, which met the inclusion criteria, were also obtained.

2.2 Adolescence and Sexuality

Adolescence is defined as the transitional period, marked by the emergence of cognitive capacities and changing societal expectations that shape and alter the self-concept. It is important to distinguish between adolescence, which refers to the psychological process of development, and puberty, the onset of sexual maturation. Adolescence, the psychological process of maturing, spans the years from 12 to 19 (for the average person).

The emerging biological, psychological and sociological changes during the period of adolescence may have implications for development and sense of self, as well as sexuality. Empirical research examining social and behavioral causes and correlates of adolescent sexuality may provide important indications as to the causes of adolescent pregnancy. However, since not all adolescents who are sexually active become pregnant, the correlates of sexual behaviors may not be exactly the same or may only partially overlap with the antecedents of risky sexual behavior and pregnancy.
2.3 Pregnancy Among Adolescents

The preponderance of empirical research on adolescent pregnancy focuses on its consequences rather than on community factors which contribute to adolescent pregnancy. Strong evidence suggests that an early transition to parenthood is associated with a wide range of negative outcomes for young mothers and their children. Adolescent childbearing also adds to the limited prospects of already disadvantaged adolescents. Compared to older mothers, the life course outcomes of teenage mothers tend to be characterized by fewer life opportunities and higher rates of psychosocial disadvantage, including: leaving school early; educational under-achievement; prolonged welfare dependence; decreased marital opportunities; less competent and more punitive parenting; maternal depression; and, greater exposure to partner violence.16,17,18

Identification of antecedent and concurrent life course factors that place young women at increased risk for early pregnancy and parenthood plays an important role in identifying key targets for future intervention and improving the success of programs. Specifically, knowledge of the community and its potential role as a risk factor is important for two reasons: 1) adverse outcomes associated with adolescent pregnancy, may in part reflect the effects of antecedent selection factors that are correlated with these problems; and, 2) an improved understanding of risk factors and life processes associated with risky adolescent sexual behavior and childbearing is necessary to improve current prevention efforts.
Antecedents of Adolescent Pregnancy

Microsystem: Individual

The self-system, or microsystem, refers to a constellation of factors, including qualities, skills, knowledge, attitudes and behaviors. Emerging sexuality accompanying adolescence may pose fundamental insecurity in the lives of young people. Insecurity includes adjusting to the altered appearance and functioning of a sexually maturing body, learning to deal with sexual desires, confronting sexual attitudes and values, experimenting with sexual behaviors, and integrating these feelings, attitudes, and experiences into a developing sense of self. Some of these psychological factors have been included in models predicting time of first intercourse, such as attitudes, values and emotional adjustment. Premarital sex may be used as an aid when developing friendships, as sex may serve as a way to combat loneliness. In addition, evidence suggests that emotional adjustment may also influence whether an individual participates in sexual activity. For example, higher levels of depression have been found to predict earlier initiation of sexual intercourse. Sexual activity may serve as a coping mechanism, or way for adolescents to feel better about themselves.

Newly discovered sexual desires and transformations are accentuated by the attention connected to being sexually attractive, the new level of physical intimacy and psychological vulnerability created by sexual encounters. These physical events in turn alter relationships between the adolescent and significant others, such as family and friends. In addition, changes are occurring within an adolescent’s cognitive and social development. In the United States, cultural proscriptions against premarital sex are counterbalanced by permissive attitudes reflected in the media and in values of many
adults. These competing perspectives co-mingle, creating a situation where adolescents are exposed to sexual material in settings of daily life, but are less able to critically evaluate it. Feelings of sexual desire and love collide with social prescriptions to show restraint, setting the stage for psychological conflict and behavioral inconsistency.

With regard to psychological factors, several prospective, longitudinal studies have found physically aggressive girls more apt than non-aggressive girls to become adolescent mothers. One study revealed a clear link between conduct problems at age eight and later pregnancy risk, with girls in the most disturbed 10 percent of the cohort being over five times more likely to become pregnant by age 18 than girls in the least disturbed 50 percent of the cohort.

Substance abuse has been documented as a contributing factor to sexual risk-taking, whereby substance use impairs individual judgment and decision-making and increases an adolescent’s risk for a pregnancy. Rates of adolescent pregnancy are elevated among licit and illicit drug users. A substantial proportion of young women becoming pregnant in adolescence years had a prior history of, or were engaging in delinquent and substance abuse behaviors. Both casual and chronic substance abusers were also more likely to engage in high-risk behaviors such as unprotected sex when under the influence of drugs or alcohol.

Strong evidence suggests cognitive deficiencies, such as lower IQ scores, poor reading, and lower educational achievement may be related to adolescent motherhood. Girls who perform poorly in school, have lower educational abilities, aspirations and motivation are more likely to become pregnant at an earlier age. One
study found educational attainment as having a delaying effect on parenthood, suggesting that females with good academic or career prospects were more likely to avoid an early pregnancy perhaps because of the greater personal costs they perceived to be associated with early motherhood.\textsuperscript{37}

A broad spectrum of individual characteristics is linked to adolescent sexual risk behaviors. Age at first intercourse is correlated with many of these risk-taking behaviors and can be used as a marker for risky sexual behavior.\textsuperscript{38} Teenagers with early onset of sexual activity tend to have more recent partners and more lifetime partners, and are less likely to use condoms, than those with later onset.\textsuperscript{39} Moreover, early age at first intercourse is independently associated with a positive sexually transmitted disease history among sexually active females.\textsuperscript{38} Proclivity toward risk-taking attitudes has also been associated with higher rates of sexual risk-taking behaviors.\textsuperscript{40} Perceived threat has been conceptualized as a product of perceived risk and perceived severity relative to a pregnancy event.\textsuperscript{41}

Differences in attitudes toward early pregnancy and single parenthood varies across racial groups in the United States.\textsuperscript{13} African American adolescents were found to participate in sexual intercourse at younger ages, at a higher percentage rate, and with more sexual partners than Caucasian adolescents.\textsuperscript{42,43} A decrease in the likelihood of condom use among Caucasian adolescents was documented, as was an increase in the likelihood of condom use among African American adolescents.\textsuperscript{44} However, contradictory results emerged when evaluating the utility of cultural and structural perspectives in accounting for interracial patterns in adolescents’ judgments about sexual intimacy.\textsuperscript{45}
The Mesosystem: Family

Family is regarded as the fundamental social system for nurturing child development. Families provide role models, shape sexual attitudes, set standards for sexual conduct, control and monitor adolescents' behaviors, and constitute the most proximate social and economic environments for adolescent development.46

Parents have many opportunities to shape the various social contexts in which their children interact; through the selection of neighborhoods in which their families reside, schools their children attend, and encouragement of extracurricular activities in which their children participate, parents indirectly shape their children’s social networks.47,48,49 Several studies have shown that living with one’s parents is often a protective factor against risky sexual behavior.50,51 Being in a dual parent family has been associated with later age of first intercourse and lower levels of sexual activity.52,53 Adolescents living with both parents implies the availability of support, supervision, and behavior control in the lives of adolescents, leaving them less likely than those in other family situations to engage in sexual risk taking behaviors.54 Baumeister et al., examined familial characteristics of Latina adolescents in two groups, one never pregnant and another pregnant or parenting, and found that those living with an intact family significantly differed from those living with one parent.55 Similarly, Capaldi et al., also found that adolescents raised in family environments characterized by single motherhood and/or multiple transitions as a result of marital breakdown have been shown to initiate sexual intercourse earlier and at an increased risk of adolescent pregnancy.56
Parenting behavior has been identified as an important source of influence on adolescent sexual activity and pregnancy. Throughout the socialization process, parents transmit their own standards of conduct, both directly through their parenting practices and indirectly through their own observable behavior. Adolescents with a high sexual risk (i.e., multiple partners, inconsistent contraception), are found to be less likely to perceive positive levels of parental support in comparison to low-risk adolescents.\textsuperscript{57} Similarly, investigations have also linked parenting behaviors with deviant peer affiliations during childhood and adolescence, such that more involved parents have children and adolescents with fewer deviant peer relationships.\textsuperscript{58,59,60}

Scaramella found that the relationship between parental warmth and pregnancy by the 12\textsuperscript{th} grade was mediated by involvement in other risk behaviors in middle adolescence (increased risk of pregnancy) and academic competence in early adolescence (decreased risk of pregnancy).\textsuperscript{61} A lack of parental support is related to depression for teen males and females.\textsuperscript{17} Similarly, this lack of support was also associated with a greater propensity for alcohol use. Adolescents who viewed their parents as being unsupportive were likely to report depressed moods and use of alcohol, but depression influenced sexual behavior for the daughters, while alcohol use was more strongly related to the sexual behavior of sons.\textsuperscript{17} In longitudinal analyses with the same sample, the relationship of parental warmth/support and adolescents’ depression was found to be linked with daughters’ sexually permissive attitudes and association with sexually active friends 1 year later.\textsuperscript{17}

Parental monitoring and supervision of adolescents’ social activities has frequently been shown to reduce risky adolescent sexual behavior. Research supports
that parents exert an influence on their adolescents’ social relationships, even when adolescents become more autonomous and independent with age. Parents maintain a greater influence on the lives of their adolescent children than their peers. Moreover, parents continue to have an influence on their adolescents’ peer group affiliations from early to late adolescence, such that parents often encourage adolescents to join one peer group over another. Greater parental monitoring is associated with less sexual activity among 9-15 year old minority youth. Other studies found that lower levels of monitoring have been associated with a greater number of sexual partners and inconsistent use of contraception.

Associations between adolescent-parent communication and adolescent pregnancy risk have been investigated in several studies. A common finding is that open, positive and frequent adolescent-parent communication about sex is associated with adolescents postponing sexual intercourse, abstaining, or having fewer sexual partners. Some investigators also reported that adolescent-parent communication is positively associated with sexually active adolescents being more likely to use effective contraception. Similarly, other studies have noted that parental communication about sexual issues prior to initial sexual intercourse and satisfaction with the quality of parent-child communication about sex are associated with adolescents’ safe sex practices. Childhood exposure to coercive child-rearing practices and dysfunctional family relationships may encourage an early pregnancy and premature transition to parenthood.

A history of sexual abuse or physical or emotional maltreatment often precedes adolescent pregnancy. Adolescents with a history of abuse are nearly three times
more likely than peers who were not abused to try to conceive, suggesting that early pregnancy may sometimes be intentional. 76 Sexually abused adolescents are more likely to have older boyfriends who were encouraging them to conceive. 76

Religious beliefs and attendance are often fostered within a family environment. 77,78 Findings on childhood religious attendance have been mixed, with some literature finding that religious attendance is protective against early sexual activity, and others finding that once a female is sexually experienced, religious belief does little to protect against a teen or non-marital birth. 79 Other findings document that religiosity and participation in religious activities are protective against adolescent pregnancy. 80

The Exosystem: Community

Literature examining community context and adolescent sexual behavior suggests that certain environmental factors may have an equal if not greater effect than individual factors on adolescents’ sexual risk behaviors. 82,85,86 Pervasive environmental pressures can reinforce adolescent sexual behavior as critical influences wane. 81 Rates of pregnancy or childbearing have been found to be related to such factors as the level of unemployment, community income, opportunities for a future, measures of community stress, and the crime rate. 82,83,84,85, 86 Overall, these studies conclude that a paucity of economic resources, racial segregation, and social disorganization seem to provide adolescents with little motivation to avoid behaviors with potentially deleterious consequences, such as pregnancy. 87

Rates of unintended pregnancy among adolescents differ according to socioeconomic factors and by racial and ethnic groups. Research examining community
effects finds that the social and economic composition of communities is an important influence on sexually related outcomes. Crane used a contextual census tract dataset that linked individuals to their neighborhood and found that African-American and white teenagers living in the most economically disadvantaged urban neighborhoods were significantly more likely to experience negative outcomes such as dropping out of high school and pregnancy, than those teens living in economically advantaged neighborhoods. Brewster researched the intersection of race, gender, and class for adolescent females’ sexual behavior with the neighborhood as the main context. This research supports the notion that neighborhood characteristics can raise structural barriers and affect development by not providing sufficient resources to adolescents. For example, Brewster found that regardless of race and ethnicity, adolescent females responded similarly to structural constraints and lack of opportunities in their neighborhoods. This research also demonstrates that, due to the ubiquitous nature of racial segregation in neighborhoods located in the United States, race can sometimes appear to be the reason for the disproportionate numbers of African-American female teenagers becoming pregnant unintentionally, at a higher rate than their white counterparts.

Studies also suggest that teens living in communities with large black or foreign-born populations are more likely to delay first intercourse, possibly due to a reduced pool of available partners of the same ethnicity or race in these communities. Adolescents who live in communities with large Hispanic populations tend to have fewer sexual partners. However, teens in these communities appear to be less likely to use effective contraception than teens from other neighborhoods.
Few studies examine a wide scope of community characteristics, such as presented in this study. Rather than observe one particular finite characteristic of community risk, this study investigates the complexities of different community determinants on an adolescent’s risk for pregnancy, allowing for a more comprehensive measure of community. If the distribution of the predictor variables is related to relative deprivation within communities where clusters of communities with a high prevalence of adolescent pregnancy are identified, then perhaps this tells us something fundamental about the impact of these characteristics on adolescent pregnancy prevention efforts.

The Macrosystem: Culture and Society

The cultural and societal values that influence individuals are found in the macrosystem. Characteristics of the macrosystem, like the exosystem, are rarely empirically studied. The most distal influence of adolescents’ risky sexual behavior is the society in which all of the other influences are embedded. By society, we mean cultural norms and traditions, large-scale policies and laws, economic conditions, and the political climate.

One specific macrosystem characteristic which may play a role in shaping cultural norms and traditions and also, through agenda-setting, can influence policy and law, is the media. Whether it is the Internet, movies, television, music videos, or books, research demonstrates that media can play a role in socialization including the socialization of adolescents, consequently even impacting their sexual risk and protective behavior. Evaluation of a nationwide mass media campaign implemented in Norway indicated that the campaign was effective in changing attitudes and practices relevant to safer sexual behaviors. In addition, trend analyses of ongoing evaluations of media-
based interventions in Switzerland have observed marked reductions in sexual risk behaviors.\textsuperscript{93} Martino et al., examined the processes mediating the relationship between exposure to television’s sexual content and adolescent’s sexual behavior. Authors found support for a model in which the relationship between exposure to television’s sexual content is mediated by safe-sex self-efficacy among African Americans and Caucasians, but not Hispanics.\textsuperscript{94}

Although adolescent childbearing can make life more complicated for young women, it is possible that adolescent childbearing is actually advantageous for some. In evaluating data from National Linked Birth/Death Files from 1983, Arline Geronimus showed that 1) black and Puerto Rican women experience much higher neonatal mortality rates at first births than do white or Mexican American women; 2) the maternal age distributions associated with neonatal mortality rates vary by race/ethnicity; 3) for white women, the highest risk of neonatal mortality is in the teens and the lowest risk is in the middle to late 20s; 4) black and Puerto Rican women in their late teens have lower neonatal mortality rates than those in their 20s and beyond; 5) for Mexican women, the 30s are higher risk than the teen years.\textsuperscript{95} For all ethnic groups, first births were most frequent at the maternal age of lowest risk for neonatal mortality. Geronimus postulated a \textit{Weathering Hypothesis} that suggests socioeconomically disadvantaged women may be subjected to many sets of health risks, the consequences of which accumulate with age. This accelerated health deterioration may make it functional for these young women to have their children early in their life course, while they are still healthy and have the maximal social and economic support from their family of origin. Despite controversy about this thesis, it is not difficult to see that early childbearing could be very important
for a group that is threatened by extreme poverty and early mortality due to poor health and violence.
CHAPTER III

RESEARCH DESIGN AND METHODOLOGY

To identify the need for targeted sexual health programming for youth in New York State, this study analyzes the distribution of adolescent pregnancy prevalence within New York State geographic locations, and, potential community factors that contribute to adolescent pregnancy. The phases of this research are designed to address the previously mentioned specific aims of this study. A mixed method design provides information from spatial and quantitative perspectives.

3.1. Geographic Unit of Analysis

ZIP codes originated as a way of classifying street segments, address ranges, and delivery points to expedite the delivery of mail. Given that ZIP codes can be associated with most places of human habitation in the United States, they present an alternative means of collecting, visualizing and analyzing spatial information. More recently, U.S. Census Bureau has produced its own ZIP code topology for area based representation, ZIP Code Tabulation Areas (ZCTAs). ZCTAs were developed as spatial units for the 2000 decennial census, as a result of requests from data users for statistical data by ZIP code area. As will be described in detail later in the dissertation, both ZIP codes and ZCTAs were utilized to define a community.

While the use of ZIP codes and ZCTAs for the spatial analysis of data may present challenges to researchers, bias related to spatial mismatch in this study was
limited for two reasons: 1) few updates have been made to New York State ZIP codes over the years; and, 2) the ZIP codes within New York State were closely aligned with the ZCTAs.

New York State ZIP codes were used to identify the prevalence of adolescent pregnancy throughout the state. U.S. Census ZCTAs were used to determine population and community factors. Henceforth I use the term ZIP code to refer to both ZIP code and ZCTA for simplicity. For the spatial analysis, 1,593 ZIP codes were included. Twenty two ZIP codes were excluded from analysis after determining specific ZIP codes were encompassed by airports, malls, or business districts. Each of the excluded ZIP codes did not contain residents. For the community regression analysis, 1,506 ZIP codes were included. Eighty nine ZIP codes were excluded from this analysis as the population of the ZIP code was less than 250 residents. The ZIP codes were excluded to avoid instability, allowing trends over time and between geographic areas to be evaluated with reasonable confidence.

County level data was not suitable for this study as it was not specific enough for local communities, especially those with large geographical areas and culturally diverse populations. For example, New York State has six counties with over one million residents, the largest being Kings County (Brooklyn) with roughly 2.5 million residents. Using county level data for this study would mask potential within county differences, making it difficult to identify issues specific to community subgroups.
3.2. Case Finding and Data Collection Methods

Adolescent pregnancy cases were identified from New York State and New York City vital records. Cases included pregnancies that resulted in a live birth, still birth or fetal death. An adolescent may be repeated in the database over time as the concern is with total adolescent pregnancies, not just one event. New York State consists of two registration areas, New York City and New York State Exclusive of New York City (also referred as Rest of State). New York City includes the five counties of Bronx, Kings (Brooklyn), New York (Manhattan), Queens and Richmond (Staten Island); the remaining 57 counties contained New York State exclusive of New York City. The New York State Department of Health processes data from live birth, death, and fetal death certificates recorded in New York State Exclusive of New York City. Through a cooperative agreement, the New York State Department of Health receives data on live births, deaths, and fetal deaths recorded in New York City from the New York City Department of Health and Mental Hygiene and on live births and deaths recorded outside of New York State to residents of New York State from other states and Canada. New York State vital statistics data is extremely valuable as the data represents the universe of events, not a sample. Vital event reporting is mandatory, as it is with other reportable diseases and conditions. However, unlike some other reportable events, reporting is virtually complete for births and deaths as families need certified copies for their legal, administrative, financial and governmental purposes. In addition, vital records and statistics staff see the records as they come in, presenting opportunities to evaluate data quality and to look for systematic errors causes by specific hospitals, and birthing centers.
The vital records system was essential to this study, allowing for enhanced speed of data acquisition and processing and accuracy of information.

New York State Public Health Law (4130) defines a live birth as the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which, after such separation, breathes or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached. Fetal deaths include both spontaneous fetal deaths and induced abortions. Fetal death is defined by New York State Public Health Law (4160) to be the death prior to the complete expulsion or extraction from its mother of a product of conception; the death is indicated by the fact that after such separation, the fetus does not breathe or show any other evidence of life such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles. New York State Public Health Law requires the registration of all fetal deaths regardless of the gestation of the fetus. Unlike birth and death registration, the registration of fetal mortality is not uniform across the United States. Categories used to report operative procedure for induced abortions include dilation and curettage, suction and curettage, dilation and evacuation, saline injection, prostaglandin injection, medical (non-surgical) and other that includes hysterectomy, hysterectomy and other procedures. Beginning in 1998, New York City combined saline and prostaglandin injection and this combination is reported in the saline injection procedure.
While the U.S. Census data is used for a myriad of research objectives, this public data was used here to assist in investigating trends in adolescent childbearing, determine community risk factors for adolescent pregnancy, and provide targeted information for pregnancy prevention program planning and funding allocations. The age-specific population of each ZIP code was determined using yearly updated population estimates of the 2000 U.S. Census. For the community regression analysis, indicators from the 2000 U.S. Census were also utilized.

In 1990 and 2000, the U.S. Census Bureau used two questionnaires to collect data: 1) Short Form Questionnaire; and, 2) Long Form Questionnaire. The Short Form Questionnaire contains 100-percent data and was the information compiled from the questions asked of all people and about every housing unit. Population items included sex, age, race, Hispanic or Latino, household relationship, and group quarters. Housing items included occupancy status, vacancy status and tenure (owner occupied or renter occupied). The Long Form Questionnaire polls a sample of households, resulting in data sets that are statistically weighted to represent the entire population. In 2000, the majority of households received a short-form questionnaire asking a minimum number of questions. A representative sample of households received a long-form questionnaire.

Data collected from the census Short Form Questionnaire is presented in Summary Files 1 and 2 (SF1 and SF2); Summary Files 3 and 4 (SF3 and SF4) were compiled from the information obtained through the Long Form Questionnaire. Summary File 1 and SF 3 files were released as individual files for each of the 50 states,
the District of Columbia, and Puerto Rico; and for the United States. SF1 provided numbers on the exact data collected, and included information compiled from questions asked of all people about every housing unit. Summary File 3 contains the sample data, which is the information compiled from the questions asked of a sample of all people and housing units. The SF1 and SF3 data from the 2000 U.S. Census was used in determining the community model developed in this research.

*American Community Survey*

The American Community Survey (ACS) 5-year 2006-2010 population estimates were used to provide timely demographic data by geographic location. The demographic variables obtained from the ACS were comparable to those selected from the U.S. Census 2000.

In 2005, the U.S. Census Bureau launched the American Community Survey in three million households nationwide. The ACS is a monthly survey designed to collect community information on key demographic, social, and economic characteristics of America’s population and housing. The ACS samples roughly 250,000 housing units across the country each month, or 3 million each year. Over a five-year period, the ACS collects data on 15 million addresses, enough to produce estimates for areas as small as census tracts and block groups. By comparison, the 2000 Census long form, which also produced data for census tracts and block groups was sent to 19.2 million addresses.

Testing of the survey’s concept began on a small scale in 1996, in four counties. Over time, the Census Bureau added more test sites, eventually conducting the survey in 36 counties around the country. Starting in 2000, the Census Bureau also fielded a national sample of roughly 700,000 households, call the Census 2000 Supplementary
Survey (C2SS). The Supplementary Survey continued in the field in subsequent years (2001-2004). Data from the 36 test sites and Supplementary Survey provide the basis for evaluations of the ACS.

Unlike the census, the ACS is not designed to provide total population counts. The census, which reaches every home and group facility in the United States, counts the number of people and produces official population numbers for all governmental units (nation, states, counties, cities, townships, census tracts and blocks). By contrast, the primary purpose of the ACS, which collects information from only a sample of the population, is to produce profiles of key population and housing characteristics for state and local communities. The decennial census (and population estimates issued annually during the decade) will continue to be the official sources of population counts for apportionment, legislative redistricting, and the allocation of federal program funds based solely on population. In contrast, the ACS will provide updated information on important characteristics of America’s population, including income, employment and occupation, education, primary language, disabilities, marital status, commuting patterns, and housing. The ACS can also be used to track changes among key population subgroups such as children, the elderly, racial and ethnic minorities, and immigrants.

The U.S. Census Bureau releases data from the ACS in the form of both single-year and multiyear estimates. In 2010, the ACS provided the first 5-year estimates of demographic, housing, social, and economic characteristics for the nation, states, cities, counties, and other small geographic areas. From here forward, these 5-year estimates will be updated annually by removing the earliest year and replacing it with the latest one, and will provide, for the first time, the ability to monitor social and economic trends
in local communities. Additionally, these estimates represent concepts that are fundamentally different from those associated with sample data from the decennial census long form.

The results of the 5-year 2006-2010 ACS, were intended to provide demographics down to the ZIP code level for the first time. Unfortunately, the Census Bureau did not provide ZIP code data with the most current data release, and do not plan on including it until the 2011 estimates, at least until late 2012 – 2013. As a result, the independent variables used in the analysis were matched to those used in the 2000 U.S. Census, downloaded at the census tract level. For the community regression analysis, demographic variables downloaded from the ACS at the census tract level, were matched to those variables chosen from the 2000 U.S. Census.

As this study was conducted, demographic data for New York State were not yet released from the 2010 U.S. Census. In addition, the 2010 ACS estimates were not released at the ZIP code level. Since ZIP codes and census tracts have been developed independently, there is little correspondence between the boundaries of the two types of units. However, the U.S. Census created a relationship file to provide associations between the 2010 ZIP codes and other 2010 Census tabulation geography: counties, county subdivisions, places (incorporated places and census designated places), census tracts, metropolitan and micropolitan statistical areas, New England city and town areas and congressional districts. The ZIP code relationship file was a nation-based comma-delimited ASCII files and include all 50 states, the District of Columbia, and Puerto Rico. This file was used to convert the ACS census tract estimates, to ZIP code estimates, for New York State using an area-weighted averaging technique. As there are roughly 5,000
census tracts and 1,500 ZIP codes in New York State, on average about three census tracts were combined into each ZIP code, though the actual number varied widely.

### 3.3. Variables

**Dependent Variables**

The dependent variable is adolescent pregnancy. This study focused on all adolescent females from 15-19 years. The time period of interest was 2006 through 2008. The time period was chosen as it includes several years of data to smooth out annual fluctuations.

**Independent Variables**

The variables included in the analysis characterized socioeconomic status, race, family and migration. These variables include: (1) percentage of households with income below the federal poverty level; 2) percentage of population older than age 24 with a Bachelor’s or higher degree; 3) percentage of single-parent households out of households with at least one child less than 18 years old; 4) percentage of total population Black alone; and, 5) percentage of total population that is Hispanic, regardless of race.

The independent measures were investigated in both the spatial analysis (5 year 2010 estimates from American Community Survey) and community analysis (2000 U.S. Census and 2010 ACS). The indicators investigated included demographic variables that were suggested by a review of the literature, the conceptual framework and available from the 2000 U.S. Census and American Community Survey. To reveal relationships or associations between the suggested quantitative variables, scatter plots were created using
Microsoft Excel. The horizontal x-axis represented the value of the chosen independent variables and the dependent variables were plotted on the vertical y-axis. After determining the extent of the relationship between these exploratory variables, a principal components analysis approach was conducted. Principal component analysis was chosen to better understand the structure of the set of variables, to identify clusters of variables that measured the latent constructs, and reduce the number of variables in the scale if there were collinear items. The main purpose of the principal component analysis was to reduce the complexity of the data by reducing the number of variables contained in the study without losing a significant amount of variance.
CHAPTER IV

DATA ANALYSIS

4.1. Spatial Analysis – Aim One and Aim Two

The purpose of conducting a spatial analysis was to identify statistically significant clusters of communities (represented by ZIP codes) with a high prevalence of adolescent pregnancy that either persist or disappear after adjustment for race, ethnicity, education, single parent household, and poverty. The 2010 ACS estimates were used to determine community factors as this is most timely and relevant data available for New York State.

For this study, a cluster of communities with a high prevalence of adolescent pregnancy is defined as the occurrence of a greater than expected number of cases of adolescent pregnancy within a group of adolescents. Identification of space variations in prevalence patterns can provide important clues for further in-depth study of the presence and control of adolescent pregnancy. Spatial clustering is defined as a general irregular spatial distribution of cases that is not confined to one particular small area.

Clusters have been analyzed previous using several statistical and epidemiological approaches. In this study, the software package SaTScan was used. SaTScan uses aspatial scan statistic to identify and test for the significance of adolescent pregnancy clusters. Although several variants of the spatial scan statistic are available, Kulldorf’s spatial scan statistic was chosen for four reasons. First, it adjusts for both heterogeneous population density and for any number of confounding variables. Second, pre-selection bias is improved upon through the search of clusters without specifying size or location.
Third, the likelihood ratio based test statistic takes multiple testing into account and
delivers a single p value for the test of the null hypothesis. Lastly, if the null hypothesis
is rejected, the approximate location of the cluster that caused the rejection can be
specified.

Kuldorff’s SaTScan, the software that performs the test for various retrospective
and prospective cluster detection and surveillance, is widely used by federal and state
agencies, such as the National Cancer Institute and the New York State Department of
Health. SaTScan can work by any of the following methods: a) evaluate reported spatial
or space-time disease clusters and determine if they are statistically significant, b) test
whether a disease is randomly distributed over space, time, or space-time, c) perform
geographical surveillance of disease, d) test geographical areas of significantly high or
low rates, and e) perform repeated time-periodic disease surveillance for the early
detection of disease outbreaks. The SaTScan software is described in detail by Kulldorff
and colleagues in a series of papers. It is summarized here to provide context for
the study approach and to introduce the interaction with and interpretation of SaTScan
results. For adolescent pregnancy, the null hypothesis of the Kuldorff spatial scan
statistic is that adolescent pregnancy cases are randomly distributed in geographic space
and the expected case count is proportional to the population at risk, adjusted by age or
other risk factors. The alternative hypothesis is that there are increased cases within an
area as compared to the outside areas. The spatial scan statistic, in determining a
significant cluster, operates by imposing a circular window on the map, whose circle
center moves over each point location so that the window includes different sets of
neighboring points at different positions. At each point location, the radius of the
circle is increased continuously from zero to a user-defined maximum radius, also referred to as the maximum size.\textsuperscript{102} The maximum size parameter sets an upper bound on the circle radius by specifying the maximum percentage of the total population at risk within the circle. A circle scanning window was chosen because this shape has been shown to be effective at highlighting general areas or regions of concern. The default maximum size value, as recommended by the SaTScan user guide, is set to 50 percent of the total population at risk.\textsuperscript{102} Therefore, a reported cluster can contain at most 50 percent of the total population at risk.

For each circle a likelihood ratio statistic is computed based on the number of observed and expected cases within and outside the circle and compared with the likelihood under the null hypothesis. To detect potential clusters, SaTScan calculates a likelihood ratio for each circle. The likelihood ratio for each circle is equal to:

\[
\left( \frac{\frac{c}{E[c]}}{\frac{C-c}{C-\hat{E}[c]}} \right)^{C-c} I( )
\]

where \( C \) is the total number of cases, \( c \) is the observed number of cases within a circle, and \( E[c] \) is the adjusted expected number of cases within the circle. \( I( ) \) is the binary indicator that facilitates identification of high risk clusters. For the purpose of this analysis, SaTScan was set to scan for high risk clusters; therefore, \( I( ) \) was set to equal ‘1’ when \( c > \hat{e} \) and equal to zero otherwise.\textsuperscript{102} The likelihood function is maximized over all windows, identifying the window that constitutes the most likely cluster. This is the cluster least likely to have occurred by chance. The likelihood ratio for this window is noted and constitutes the maximum likelihood ratio test statistic. An associated p-value, based on Monte Carlo simulations, is computed and used to evaluate whether the cases
are randomly distributed in space or not. For each simulation the likelihood ratio statistic is computed and the actual value is compared with the set of simulated values to find the significance probability. Secondary clusters, also identified by SaTScan, have a significantly large likelihood ratio, but are not the primary cluster. Secondary clusters occur as a result of slight alterations to the circle radius or relocation of the circle center to a nearby location, adding or removing locations to the circle.

The next step was to determine if the statistically significant clusters would differ when the model was adjusted for the chosen risk factors: race (Black alone), ethnicity (Hispanic), family structure (single parent household, education (Bachelor’s degree or higher), poverty (below federal poverty level) and all combined. While SaTScan is able to adjust for categorical risk factors, SaTScan cannot in itself do an adjustment for continuous risk factors. Therefore, the risk factor adjusted expected number of cases for each community were calculated using regression analysis in SAS 9.1.1. The expected cases then replaced the raw population numbers in the population file and were run in SaTScan. This process was conducted for each risk factor, producing different sets of expected case counts, which were run separately. By adding one risk factor at a time, and identifying the change, the underlying geographic distribution of that risk factor can be observed, as well as the how that risk factor affects the distribution of cases in New York State. Underlying geographic distribution of a combination of all risk factors was also identified.

At this time, SaTScan does not output a graphic file that can be immediately read into a geographic information system. Output from SaTScan was used as an input dataset in Python for mapping into Google Earth. Python is a general purpose programming
language used for processing text, numbers and images. Existing code developed by the New York State Cancer Registry was modified to convert the SaTScan text output into a kml file, which automatically opened and displayed clusters in Google Earth.\textsuperscript{103}

4.2 Community Analysis Modeling Framework – Aim Three

To determine risk factors associated with clusters of communities with a high prevalence of adolescent pregnancy a generalized linear mixed model (GLMM) was employed. The model assumes a continuous outcome variable which is linearly related to a set of explanatory variables. This modeling provides estimates of the unexplained variance in the outcome that is due to unobserved community factors, generally known as the random effect.\textsuperscript{104} Accordingly, the model includes a random intercept at the cluster level to capture heterogeneity among clusters. The model assumes that within each cluster, contraceptive use follows a binomial distribution. Further, the cluster means themselves vary according to a normal distribution with a mean (random intercept) and a variance. In other words, they assume that the average in each community differs on the basis of community characteristics. The variation in random intercepts is captured as the random effect. A significant random effect indicates unexplained heterogeneity between clusters after other factors in the model.

The mixed model is equal to $Y=\alpha+\beta X+\gamma Z+\varepsilon$ where $Y$ is the vector of observed response data values; $\alpha$ is the population intercept; $\beta$ is the vector of unknown fixed effects parameters; $X$ is the known design matrix for the fixed effects; $\gamma$ is the vector of unknown random effects parameters; $Z$ is the known design matrix for the random effects and $\varepsilon$ is the vector of random errors.
For this study, the model is as follows, \( i = 1, \ldots, n \) ZIP codes, \( y_i \) is the observed count of adolescent pregnancy cases in each ZIP code. For every observational unit (ZIP code), there is a response variable, adolescent pregnancy counts, for the time period 2006 - 2008. The population at risk, \( n_i \), includes all females in ZIP code \( i \), between the ages 15-19, between 2006 - 2008. The community predictors, \( X_1, \ldots, X_p \), include race (Black alone), ethnicity (Hispanic), family (single parent), poverty status, and education (at least a Bachelor’s degree). The linear coefficients, \( \beta_1, \ldots, \beta_p \), relate the degree and strength of association between the outcome and community predictor. As a count of a rare occurrence, \( y_i \) was considered to be distributed as a generalized negative binomial random variable, where the mean count for the distribution is modeled via a canonical log link to a linear function of the predictors to estimate relative risk among the categories of each predictor.

The effect of spatial location (L) was also evaluated to account for possible unidentified spatially varying covariables in the model. For this, a random effect was added to the county which exhibited the largest part of a ZIP code. The result was a model with a random addition to the intercept. This process reduced the residual spatial autocorrelation and potential confounding and provided a smoothing effect, in that the predicted cases is adjusted toward a common local value. The actual observed cases could be compared to what is expected, given a statewide reference model. The expected value of \( y_i \) given \( \{X_1, \ldots, X_p, L\}_i \) = \( E[y_i \mid \{X_1, \ldots, X_p, L\}_i] = n_i \exp(\beta_1 X_1 + \beta_p X_p + L_i) \). The probability of observing \( y_i \) cases in a cell defined by the \( i \)th ZIP code is:

\[
P(Y_i = y_i, \mu_i, \lambda_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha + 1) \Gamma(\alpha)}
\]
For $\alpha \geq 1$ and $\Gamma$ is the gamma function, where the mean for each cell is linked to the covariables by the log link, after offsetting by the population within each cell, $n_i$, such that $\mu_i = n_i \exp(X_i^T \beta + \lambda_c)$ for risk factor vector $X_i$ linear coefficient vector $\beta$ and county specific random effect $\lambda_c$, which is assumed to be iid $N(0, \sigma^2_c)$ for inter-county variance $\sigma^2_c$. The term $\alpha$ is a scale parameter that is estimated from the data and increases as the variance exceeds the mean.

The end result of the above modeling is a community risk adjusted estimate of communities with respect to adolescent pregnancy. The risk adjusted estimates were then run through a SAS program to determine the three cutoff points. The risk adjusted estimates were then categorized into tertiles, based on these cutoff points. Risk adjusted estimates were then categorized into low, middle or high level of risk. This approach avoids the need to ensure a linear relationship between risk factor and the response, while also allowing presentation of the results as risk of a medium or low category relative to a high category.

Modeling was carried out with the SAS 9.3 GLIMMIX procedure. The GLIMMIX procedure performs an estimation and statistical inference for generalized linear mixed models. The model fit by the GLIMMIX procedure extended the generalized linear model by incorporating correlations among the responses. This was accomplished using random effects in the linear predictor and modeling the correlations among the data directly. For this study, the default approach of pseudo-likelihood was used. Overall model fit was diagnosed by comparing the generalized chi-square statistic to the degrees of freedom, which are equivalent when the model is properly specified. Observed values were graphed and regressed against the model predicted
values to assess model predictability. Since the observed number of cases depends on varying underlying population sizes, Pearson-type residuals were analyzed because they are standardized by the square root of the variance of their respective observations. Residuals were assessed for outliers and independence with respect to the model predicted values, and were also tested for spatial autocorrelation by Moran’s I statistic\textsuperscript{107} to evaluate the impact of adding a spatial random effect.

4.3 Geographical Mapping

Maps can be an effective method for presenting data that varies geographically. Maps provide a spatial picture of the data, and allow clusters or areas of concentration to be easily identified. Spatial data refers to anything that can be referenced based on its physical location, such as ZIP codes.

To demonstrate ZIP codes with high cases of adolescent pregnancy, thematic mapping was used. SAS 9.3 statistical software was used to map adolescent pregnancies, one of the procedures available with SAS/GRAPH. Maps were produced using a combination of a map data set and response data set. The map data set contained information necessary to draw the map boundaries, and the response data set contained code specific adolescent pregnancies. Observations between the response data set and the map data set were matched, resulting in a geographical display of the response data set. A chloropleth map, using colors to distinguish map areas was used for this study. Maps provide a visual representation of New York State ZIP codes and illustrate the spatial distribution of adolescent pregnancies in this study.
5.1 Descriptive Analysis

The adolescent pregnancy dataset represented females, ages 15-19 years, who became pregnant from 2006-2008 and reside in New York State. During 2006-2008, roughly seven percent of the average of 751,499 births occurring each year in New York was to an adolescent. The average yearly rate of adolescent pregnancy for New York State was 58.4 per 1,000 for the female population aged 15-19 years. A map of the corresponding 2006-2008 adolescent pregnancy rates stratified by ZIP code is shown in Figure 5.1. This map clearly indicates the nonhomogeneous distribution of adolescent pregnancy over the geographic regions of the state. A total of 37,928 adolescent pregnancies from 1593 communities were identified and included in the analysis. Total adolescent pregnancies ranged from 0 to 572, with a mean of 24 cases. The total female population within New York State communities ranged from 250 to 648,916.
Figure 5.1. Adolescent pregnancy rates, aggregated from 2006-2008
Aim One: Cluster analysis results from unadjusted model

The Poisson model was used to calculate expected clusters of high adolescent pregnancy prevalence within each community. The spatial analysis was conducted without prior assumptions of the size, location or duration of excess communities with high adolescent pregnancy prevalence. The default scanning setting of a maximum spatial cluster size of 50 percent of the population at risk was used. The 50 percent of population at risk parameter was recommended by Kulldorf as an optimal value setting that maximizes the effect of potential cluster detection. This would mean that a cluster would comprise at most 50 percent of the population at risk. Under the null hypothesis, when there are no community factors, the expected number of cases in each area is proportional to its population size.

Adjusting only for the underlying population at risk, nine statistically significant \((p < 0.05)\) communities with high adolescent pregnancy prevalence were identified in New York State (Table 5.1 and Figure 5.2). The most likely cluster was located in New York City. This cluster included 149 communities and 63 percent more cases than expected, when compared with the rest of the state. This cluster was significant \((p < .01)\).

In order to discover the possibility of smaller clusters in the study area, the same analysis was performed with a spatial cluster size of < 25 percent. This analysis identified the most likely significant cluster and secondary clusters in the same locations.
Figure 5.2. Statistically significant clusters of communities with high adolescent pregnancy prevalence.
SaTScan also identified secondary clusters in the data set in addition to the most likely cluster, which are ordered according to their likelihood ratio test statistic.

Secondary clusters are geographically concentrated in the Western region of New York; the Central region; and the Capital District region. Of these secondary clusters, eight proved to be statistically significant at the level p < .01. Secondary clusters labeled ‘2’, ‘3’, and ‘5’ (Table 5.1) are centered in the Cities of Rochester, Buffalo and Niagara Falls. Combined, these three clusters observed 2,192 adolescent pregnancies among 31 communities. Corresponding relative risks were 2.68, 1.83 and 2.25, respectively.

Secondary clusters ‘4’, ‘6’ and ‘9’ are located in the Central Region. The three clusters centered among Syracuse, Utica and Binghamton, respectively. The clusters encompassed nine communities and observed 756 adolescent pregnancies. The number of adolescent pregnancies is 176 percent (‘4’), 85 percent (‘6’) and 51 percent (‘9’) higher than expected.
The Capital District represents the last grouping of secondary clusters (‘7’ and ‘8’). The clusters are centered within the Cities of Albany and Schenectady. The clusters observed 340 adolescent pregnancies among nine communities. Within the Capital District clusters, adolescent pregnancies are 102 percent (‘7’) and 73 percent (‘8’) higher than expected for the rest of the state.

When combining all clusters, the 205 communities are responsible for 60 percent or 22,747 adolescent pregnancies, with an average rate of 5,874.9 pregnancies per 100,000 persons per year.

**Aim Two: Cluster results from adjusted analyses**

When a cluster of communities with a high prevalence of adolescent pregnancy cannot be dismissed as a chance occurrence, the question to ask is what may be the underlying causal mechanisms. A crucial next step is to look at some of the known or hypothesized community factors. In this study, five such factors are analyzed: race, education, ethnicity, poverty, and single parent households. Results were generated by the spatial scan method when the maximum cluster size was set to 50 percent of the total population.

After adjusting for education, ten statistically significant (p < .05) clusters of communities with high adolescent pregnancy prevalence were identified (Figure 5.3.). The most likely cluster was found in a similar geographic location as that of the primary cluster in the unadjusted analysis (Figure 5.2). Statistically significant secondary clusters also followed similar patterns. SaTScan identified 225 communities (14 percent) where the prevalence of adolescent pregnancy is higher than expected. The statistically
significant clusters were responsible for 67 percent or 25,255 adolescent pregnancies over the three years, with an average rate of 12,377.9 pregnancies per 100,000 persons per year.

The largest cluster was located in New York City, encompassed 178 communities and observed 104 percent more adolescent pregnancies than expected. The second largest cluster was located in Rochester. This cluster encompassed 15 communities and observed 96 percent more adolescent pregnancies than expected. The third cluster, located in Buffalo, contained 14 communities and observed 54 percent more adolescent pregnancies than expected. The remaining seven clusters were located in Syracuse (two clusters), Albany, Schenectady, Utica, Niagara Falls and Binghamton. These clusters included 18 communities and observed 41 to 119 percent more adolescent pregnancies than expected.

In this analysis, an additional cluster proved significant, not located on the unadjusted model. This cluster was located in the City of Syracuse (Figure 5.3 cluster labeled ‘7’). The cluster included two communities, with ZIP codes 13210 and 13205.
Figure 5.3. Statistically significant clusters of high adolescent pregnancy prevalence, adjusted for education
After adjusting for race, ten statistically significant (p < .05) clusters of communities with high adolescent pregnancy prevalence were identified (Figure 5.4). Analysis identified 186 communities (11 percent) where prevalence of adolescent pregnancy is higher than expected. The statistically significant clusters were responsible for 59 percent or 22,376 adolescent pregnancies over the three years, with an average rate of 5,618.3 adolescent pregnancies per 100,000 persons per year.

The largest cluster is located in New York City. This primary cluster comprised 133 communities with 121 percent more adolescent pregnancies than expected. The second largest cluster is found in Rochester. This cluster includes 11 communities with 151 percent more adolescent pregnancies than expected. These two clusters correspond with the two largest clusters identified in the previous unadjusted analysis, although notably decreased in size. The third largest cluster is located in Buffalo. These two clusters included 17 communities with 74 percent more adolescent pregnancies than expected. The remaining seven clusters were located in Syracuse (two clusters), Utica, Niagara Falls, Albany, Schenectady and Binghamton. The clusters contained 25 communities, with 56 – 137 percent more adolescent pregnancies than expected.
Figure 5.4. Statistically significant clusters of high adolescent pregnancy cases, adjusted for race
After adjusting for ethnicity, ten statistically significant (p < 0.05) clusters of communities with high adolescent pregnancy prevalence were identified (Figure 5.5). Results identified 232 communities (15 percent) in which clusters of communities with high prevalence of adolescent pregnancy were higher than expected. The statistically significant clusters are responsible for 25,812 adolescent pregnancies over the three years (68 percent), with an average rate of 5,320.5 adolescent pregnancies per 100,000 persons per year.

The largest cluster was located in New York City. The primary cluster contained 178 communities and observed 103 percent more adolescent pregnancies than expected. The second largest cluster is located on the western part of the state, and contained 11 communities with 170 percent more adolescent pregnancies than expected. The third largest cluster included 17 communities and 93 percent more adolescent pregnancies than expected. The remaining eight clusters comprised 26 communities and 63 – 164 percent more adolescent pregnancies than expected.
**Figure 5.5.** Statistically significant clusters of high adolescent pregnancy cases, adjusted for ethnicity.
After adjusting for poverty, ten statistically significant (p < 0.05) clusters were identified (Figure 5.6). SaTScan identified 212 communities (13 percent) in which clusters of communities with high prevalence of adolescent pregnancy were higher than expected. The clusters accounted for 63 percent or 23,814 adolescent pregnancies over the three years with an average rate of 5,494.5 adolescent pregnancies per 100,000 persons per year.

The largest cluster was located in New York City. This primary cluster contained 161 communities with 121 percent more adolescent pregnancies than expected. The second largest cluster was located in the Western region of New York and included 151 percent more adolescent pregnancies than expected. The third largest cluster included 17 communities and 73 percent more adolescent pregnancies than expected. The remaining six clusters comprised 22 communities and 78 – 136 percent more adolescent pregnancies than expected.
Figure 5.6. Statistically significant clusters of high adolescent pregnancy cases, adjusted for poverty.
In the cluster analysis adjusted for single parent households 13 significant clusters (p < 0.05) were detected. These clusters are represented in Figure 5.7. Analysis identified 235 communities (15 percent) in which clusters of communities with high prevalence of adolescent pregnancy were higher than expected. The significant clusters were responsible for 65 percent of or 24,609 adolescent pregnancies, with an average rate of 2,791.8 adolescent pregnancies per 100,000 persons per year.

The most likely cluster was located in New York City. The cluster comprised 159 communities and 33 percent higher adolescent pregnancies than expected. The second largest cluster is located in Rochester. This cluster included 16 communities, and 48 percent higher adolescent pregnancies than expected. The third largest cluster is located in Syracuse. This cluster contained six communities and 109 percent higher adolescent pregnancies than expected. The remaining clusters were located in Buffalo, Utica, Binghamton, Niagara Falls, Schenectady, Kingston, Jamestown, Suffolk, Albany and Watertown. These clusters comprised 54 communities, with 40 – 134 percent more adolescent pregnancies than expected.
Figure 5.7. Statistically significant clusters of high adolescent pregnancy cases, adjusted for single parent households.
Adjusting for all risk factors combined

The final spatial analysis adjusted for race, education, ethnicity, poverty and single parent households combined. When adjusting for all risk factors, no statistically significant clusters remained.

Summary of results

With the adjustments for education, race, ethnicity and poverty, the spatial analysis produced broadly similar patterns across the Western, Central, Capital District and New York City regions as that of the unadjusted analysis (Figures 5.3-5.6). An additional cluster was identified for education alone, race alone, poverty alone, and ethnicity alone. The additional cluster was located in the City of Syracuse. The additional cluster contains two communities and is roughly six miles from the Syracuse cluster identified in the unadjusted analysis. The addition of each community factor into the analysis did not remove existing clusters of communities with high adolescent pregnancy prevalence; indicating, a lack of presence of any ‘significant’ impact of these community factors.

After adjusting for single parent households, four additional clusters emerged (Figure 5.7). After adjusting for single parent households, the additional clusters were located in areas of the state not otherwise identified in the unadjusted analysis, or when adjusting for race, ethnicity, education or poverty. The cities identified included: Kingston, Jamestown, Suffolk and Watertown. When a new cluster emerges after adjusting for a specific community factor, it is because the relationship between the community factor suppressed the true relationship with adolescent pregnancy. Meaning,
the community factors adjusted for in these instances, suppressed the emergence of these communities before they were explicitly controlled for. If it were not for a favorable profile with respect to the prevalence of single parent households, specifically, these communities would be identified as having a high prevalence of adolescent pregnancy.

When adjusting for all combined community factors, all clustering of communities with high adolescent pregnancy prevalence disappeared. This suggests that clusters of communities with high adolescent pregnancy prevalence was due to a combination of high risk community factors, including, percentage of residents living in poverty, racial and ethnic minorities and percentage of single parent households.

*Tracking of Community variables by cluster*

To confirm which community factor had the greatest impact on adolescent pregnancy prevalence, several clusters were evaluated. The tracking of clusters for this study followed a similar process described by Henry et al. when evaluating geographic disparities in colorectal cancer survival.\(^{108}\) This technique, while documented in published cancer research, has not been documented as frequently with public health studies. The first cluster included part of the City of Albany. In tracking this cluster, the observed and expected cases and relative risk were identified from the risk estimates location file. SaTScan does not report results like this directly; however, the file containing this information can be generated as part of the program’s advanced output options. The results are found in Table 5.2. With each set of analyses, the observed number of adolescent pregnancies remained the same, and the expected number varied with each adjustment. Therefore, if a community factor was positively associated with
excess risk for adolescent pregnancy prevalence, it increased the expected number of such adolescent pregnancies, and decreased the observed-to-expected ratio. Factors that suppress the risk for high adolescent pregnancy prevalence, when adjusted for, reduced the number of such adolescent pregnancies expected, and increased the observed-to-expected ratio.

In comparing results of the Albany cluster, adjusting for education and ethnicity, education has a lower observed-expected ratio (1.97 vs. 2.13), thus education is seen as a more important variable. In comparing education and race, race has a lower observed-expected ratio (1.91 vs. 1.97); therefore, race would be a more important factor then education and ethnicity. When comparing single parent households, the observed-expected ratio is the lowest (1.51 vs. 1.91, 1.97). In part of the City of Albany, therefore, single parent households was considered the most important factor of the community variables analyzed. When adjusting for all factors, the observed and expected adolescent pregnancies are almost equal, explaining the variation.

Table 5.2. Tracking results for part of the City of Albany to confirm factors of most importance, after adjustments.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Expected</th>
<th>Observed:Expected Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>162</td>
<td>80</td>
<td>2.02</td>
</tr>
<tr>
<td>Race, only</td>
<td>162</td>
<td>85</td>
<td>1.91</td>
</tr>
<tr>
<td>Ethnicity, only</td>
<td>162</td>
<td>76</td>
<td>2.13</td>
</tr>
<tr>
<td>Education, only</td>
<td>162</td>
<td>82</td>
<td>1.97</td>
</tr>
<tr>
<td>Poverty, only</td>
<td>162</td>
<td>85</td>
<td>1.91</td>
</tr>
<tr>
<td>Single parent, only</td>
<td>162</td>
<td>102</td>
<td>1.5</td>
</tr>
<tr>
<td>All factors</td>
<td>162</td>
<td>161</td>
<td>1.01</td>
</tr>
</tbody>
</table>
To further understand the extent to which geographic patterns of communities with high adolescent pregnancy prevalence can be explained by community factors, part of the City of Niagara Falls was evaluated. The observed-expected ratios for each factor are found in Table 5.3. After adjusting for race and ethnicity, ethnicity has a lower observed-expected ratio (2.20 vs. 2.25), thus ethnicity is seen as a more important variable. In comparing poverty and ethnicity, poverty has a slightly lower observed-expected ratio (2.17 vs. 2.20); therefore, poverty would be a more important factor than race and ethnicity. In comparing single parent households and poverty, single parent households has a lower observed-expected ratio (1.98 vs. 2.17); therefore, single parent households would be a more important factor than poverty, ethnicity and race. In comparing education and single parent households, education has a lower observed-expected ratio (1.41 vs. 1.98); therefore, education would be a more important factor than single parent households, poverty, ethnicity and race. As a result, education was considered the most important community factor analyzed for this part of the City of Niagara Falls.

**Table 5.3.** Tracking results of part of the City of Niagara Falls to confirm factors of most importance, after adjustments.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Observed</th>
<th>Expected</th>
<th>Observed:Expected Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>158</td>
<td>70</td>
<td>2.25</td>
</tr>
<tr>
<td>Race, only</td>
<td>158</td>
<td>72</td>
<td>2.20</td>
</tr>
<tr>
<td>Ethnicity, only</td>
<td>158</td>
<td>65</td>
<td>2.44</td>
</tr>
<tr>
<td>Education, only</td>
<td>158</td>
<td>106</td>
<td>1.41</td>
</tr>
<tr>
<td>Poverty, only</td>
<td>158</td>
<td>73</td>
<td>2.17</td>
</tr>
<tr>
<td>Single parent, only</td>
<td>158</td>
<td>80</td>
<td>1.98</td>
</tr>
<tr>
<td>All factors</td>
<td>158</td>
<td>157</td>
<td>1.01</td>
</tr>
</tbody>
</table>
The third cluster of communities evaluated included part of the City of Binghamton. The observed-expected ratios for each factor are found in Table 5.4. After adjusting for education and single parent households, education demonstrated a lower observed-expected ratio (1.83 vs. 1.98), thus, education is seen as a more important variable. In comparing ethnicity and education, ethnicity has a lower observed-expected ratio (1.63 vs. 1.83); therefore, ethnicity would be a more important factor than single parent households or education. In comparing race to ethnicity, race has a lower observed-expected ratio (1.56 vs. 1.63); therefore, race is a more important factor than single parent households, education and ethnicity. When comparing poverty to race, poverty has a lower observed-expected ratio (1.48 vs. 1.56); therefore, poverty would be a more important factor than single parent households, education, ethnicity and race. As a result, poverty is considered the most important community factor analyzed for this part of the City of Binghamton.

Table 5.4. Tracking results of part of the City of Binghamton to confirm factors of most importance, after adjustments.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Expected</th>
<th>Observed:Expected Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>160</td>
<td>106</td>
<td>1.51</td>
</tr>
<tr>
<td>Race, only</td>
<td>160</td>
<td>103</td>
<td>1.56</td>
</tr>
<tr>
<td>Ethnicity, only</td>
<td>160</td>
<td>98</td>
<td>1.63</td>
</tr>
<tr>
<td>Education, only</td>
<td>160</td>
<td>88</td>
<td>1.83</td>
</tr>
<tr>
<td>Poverty, only</td>
<td>160</td>
<td>108</td>
<td>1.48</td>
</tr>
<tr>
<td>Single parent, only</td>
<td>160</td>
<td>80</td>
<td>1.98</td>
</tr>
<tr>
<td>All factors</td>
<td>160</td>
<td>158</td>
<td>1.01</td>
</tr>
</tbody>
</table>
Part of the City of Watertown was identified as an additional cluster when explicitly controlling for single parent households. As discussed previously, this is due to the suppressor effect of single parent households. To confirm the favorable profile of single parent households in this cluster of communities, the observed-to-expected ratios were calculated for all the unadjusted, as well as all adjusted factors (Table 5.5). After evaluating all community factors, controlling for single parent households resulted in the largest increase in the observed-to-expected ratio (1.45, 1.50, 1.75, 1.83 vs. 2.34). This is consistent with the emergence of this cluster, after explicating controlling for single parent households. If it were not for a favorable profile with respect to single parent households, this cluster would most likely have a high prevalence of adolescent pregnancy.

**Table 5.5.** Tracking results of part of the City of Watertown to confirm factors of most importance, after adjustments.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Expected</th>
<th>Observed:Expected Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>42</td>
<td>24</td>
<td>1.75</td>
</tr>
<tr>
<td>Race, only</td>
<td>42</td>
<td>24</td>
<td>1.75</td>
</tr>
<tr>
<td>Ethnicity, only</td>
<td>42</td>
<td>23</td>
<td>1.83</td>
</tr>
<tr>
<td>Education, only</td>
<td>42</td>
<td>29</td>
<td>1.45</td>
</tr>
<tr>
<td>Poverty, only</td>
<td>42</td>
<td>27</td>
<td>1.50</td>
</tr>
<tr>
<td>Single parent, only</td>
<td>42</td>
<td>18</td>
<td>2.34</td>
</tr>
</tbody>
</table>
5.3 Community Analysis – Aim Three

The statistical modeling framework was initially developed to analyze adolescent pregnancy as a function of select community factors obtained from the 2000 U.S. Census data. Prior to finalizing this study however, the U.S. Census released state-based 2010 estimates from the American Community Survey. To provide timely and relevant data to researchers, health professionals and pregnancy prevention planners, the ACS 2010 estimates were used to reveal the issue of adolescent pregnancy as it relates to New York State communities. While the 2000 U.S. Census data is not as current, it is used in this study as a comparison, to understand changes in the nature of the risk for adolescent pregnancy among New York State communities over the last decade. Both sets of data were incorporated into the same model and illustrated through graphical representation. Additionally, another factor, foreign born citizenship was also added to the model. This factor was chosen to identify the relative importance on adolescent pregnancy and whether it is seen as either a protective or risk factor.

As discussed in the data analysis chapter, regression analysis was used to produce an equation to determine communities with high adolescent pregnancy prevalence using several community factors. The equation has the form $Y = (\beta_1X_1 + \beta_2X_2 + \ldots + A)$ where $Y$ is the cluster of communities with high adolescent pregnancy the study intended to predict, $X_1, X_2, \ldots$, are the community factors used to determine a cluster of communities with high adolescent pregnancy prevalence, and $\beta_1, \beta_2, \ldots$, are the coefficients that describe the size of the effect the community factors are having on the cluster of communities of high adolescent pregnancy prevalence. ‘$A$’ is the value $Y$ is predicted to have when all community factors are equal to zero.
For prediction equation of Model A, which used the 2000 U.S. Census demographic data, a cluster of communities with high adolescent pregnancy prevalence = -.01872 (education) + .008477 (race) + .007101 (ethnicity) + .01619 (single parent) + .004203 (migration) + .01831 (poverty). For prediction equation of Model B, which used the ACS 2010 estimates, a cluster of communities with high adolescent pregnancy prevalence = (Model B)= -.01533 (education) + .007908 (race) + .004504 (ethnicity) + .02485 (single parent) + .007821 (migration) + .0109 (poverty). The beta coefficients comparing clusters of communities with high adolescent pregnancy prevalence and the independent variables chosen from the Census based data were mostly positive across both Model A and B (Table 5.6). The coefficients for Model A ranged from -.01872 to .01619, and for Model B, -.01533 to .02485.

As the p values are directly derived from correlations, education, race and poverty were found as the most important factors. Education, race and poverty all observed a highly significant p value of <.0001. The p value observed this same level of significance for each of the factors in both Model A and Model B. While this differs from the spatial analysis, which indicated single parent households were the most important, this is not a contradiction. In the state as a whole, the percentage of single parent households is not the most important factor; however, it may in individual communities.

The coefficient for foreign born citizens, observed in Model A was not found to be significant; however, in Model B, this was found to be significant. All other variables reveal a consistent effect from Model A to Model B with respect to direction, magnitude and statistical significance. The risk to adolescents residing in a community
increases as the following community factor increases: level of poverty among residents; number of single parent households within a community; percentage of black residents; percentage of Hispanic residents; and percentage of foreign born residents. The risk of pregnancy among adolescents decreases when the number of residents within a community that attained a Bachelor’s degree or higher increases.

Table 5.6. Beta coefficients and p values resulting from the generalized linear mixed model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>U.S. Census 2000 (Model A)</th>
<th></th>
<th>ACS 2010 (Model B)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Education (Bachelor's Degree or higher)</td>
<td>-0.01872</td>
<td>&lt;.0001</td>
<td>-0.01533</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Race (Black alone)</td>
<td>0.008477</td>
<td>&lt;.0001</td>
<td>0.007908</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Ethnicity (Hispanic)</td>
<td>0.007101</td>
<td>0.0002</td>
<td>0.004504</td>
<td>0.0005</td>
</tr>
<tr>
<td>Family Structure (Single Parent)</td>
<td>0.01619</td>
<td>0.0087</td>
<td>0.02485</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Migration Status (Foreign-born naturalized citizen)</td>
<td>0.004203</td>
<td>0.3289</td>
<td>0.007821</td>
<td>0.0095</td>
</tr>
<tr>
<td>Poverty (Below federal poverty limit)</td>
<td>0.01831</td>
<td>&lt;.0001</td>
<td>0.0109</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Summary of modeling results

Results of the generalized linear mixed model provide a community risk adjusted estimate of predicted adolescent pregnancies for each community, based on the statewide reference population. Each community risk adjusted estimate was categorized into one of four risk levels: low, middle, high or risk could not be determined. The risk adjusted estimates were calculated using the 2000 U.S. Census demographic estimates (Model A), and the 2010 ACS demographic estimates (Model B). A map of Model A and Model B is found in Figure 5.8.

Of the 1506 communities analyzed, 28 percent (419) of these communities observed an increased risk for adolescent pregnancy from 2000 to 2010 (Figure 5.9). Of these 419 communities, 55 percent observed an increase from low to medium; 40 percent observed an increase from medium to high; and, .2 percent observed an increase from low to high. Conversely, a small number of communities experienced a reduction in the risk to adolescents among their community: three percent of communities lowered risk from medium to low; and, .5 percent of the communities lowered risk from high to medium. No community lowered risk from high to low.
Figure 5.8. Comparison of adolescent pregnancy by ZIP code from 2000 – 2010.
Figure 5.9. New York State communities observing an increase in the level of risk for adolescent pregnancy, from 2000 to 2010.
The aims of this study were to: identify clusters of communities with high adolescent pregnancy prevalence; determine the association of community factors (race, ethnicity, education, poverty and single parent households), on clusters of communities with a high prevalence of adolescent pregnancy; and, determine whether the risk for adolescent pregnancy prevalence changed among communities from 2000 to 2010. A series of analyses were undertaken to: 1) test for presence of community clusters of high adolescent pregnancy prevalence, adjusted for community factors and identified their approximate location; and, 2) assess association between geographic location and community factors on adolescent pregnancy. The findings, as they pertain to each aim, are summarized below.

6.1. **Aim One:** Identify statistically significant clusters of communities with a high prevalence of adolescent pregnancy.

Adolescent prevention programs, either free-standing, or as part of a larger organization, such as a county or state health department, need to have relevant and timely information to implement effective interventions. Geographic locations at greatest risk should be targeted.

Spatial epidemiology is concerned with describing and understanding spatial variations in disease risk, using geographically indexed health and population data along
with statistical methodology and geographical information systems. Spatial epidemiology has been used to solve perplexing problems in public health for over a century. The availability of modern day statistical tools coupled with the developments in geographic information systems has resulted in major advancements in the field of public health. One of the strengths of this study was the application of the spatial statistical tool. SaTScan, applied largely to cancer studies, was adapted for use in public health assessment and program planning. SaTScan assisted in investigating adolescent pregnancy prevalence, the distribution of communities experiencing high prevalence of adolescent pregnancy and associated community factors exerting influence on this prevalence.

The SaTScan method is an increasingly popular adjunct for exploring spatial distribution of health outcomes. SaTScan is capable of analyzing spatial, temporal, and space-time data using the spatial, temporal or space-time statistics. It is designed to perform geographical surveillance of health outcomes, to detect spatial or space-time clusters of outcomes; to determine whether outcomes are randomly distributed over space, time or space and time; to evaluate statistical significance of clusters; and, to perform repeated time-periodic outcome surveillance for early detection. SaTScan has been used to confirm investigations of disease and to further explore disease patterns. This method compares all possible aggregations of neighboring populations with the rest of the study area and orders those aggregations on the likelihood that a cluster exists within the aggregated area.

The spatial scan statistic offers state and local health departments an objective way of detecting clusters of communities with high prevalence of adolescent pregnancy,
and providing information to supplement data obtained from traditional descriptive analyses. The value of using the spatial scan statistic is the ability to analyze variations in adolescent pregnancy prevalence for small and large geographical groupings, unable to be achieved by merely comparing county or regional rates. The spatial analysis can assist professionals to prioritize specific communities for intervention particularly when overall adolescent pregnancy prevalence exceeds expectation or when state or local health departments' resources are limited.

Results indicated that nine communities experience excess adolescent pregnancy prevalence that are statistically significant. The geographic locations of the clusters identified are of interest. Clusters were not found among one region or neighboring region, rather, they were spaced generously among several regions of the state. In addition, clusters were identified in urban cities, providing some evidence that initial results, based on population, might reflect geographical concentrations of adolescents who can be presumed at greater risk for adolescent pregnancy based on place of residence.

Many of the state’s cities have experienced serious population losses which have contributed to economic and fiscal decline. For example, the 2000 census documents the population of the City of Buffalo is one half that of 1950. Rochester and Syracuse have also had substantial population declines since 1950. Shifts in global manufacturing, the growth of a more technologically-based economy, and the development of suburban-based lifestyles are underlying causes of these population trends. In many parts of the state, more people now live in suburbs than in cities. The middle-class exodus over the past few decades has led to a general worsening of
socioeconomic conditions in cities.\textsuperscript{114} In general, cities tend to have greater levels of poverty, higher levels of vacant housing, a greater percentage of female headed households with children, and a higher percentage of adults with less than a high school diploma when compared to surrounding suburban neighborhoods.\textsuperscript{114} These factors reflect the effects of out-migration patterns and home-purchasing decisions, particularly where conditions within a city are significantly worse when compared to outer-fringe and suburban neighborhoods.\textsuperscript{115}

Given these dynamics, statistically significant clusters of communities with high adolescent pregnancy prevalence could be postulated to result from negative effects of urban features and larger, denser and more heterogeneous cities, which increase the potential for negative outcomes, such as adolescent pregnancy.\textsuperscript{116} The urbanism theory suggests that large, dense, and diverse places foster social disorganization and alienation among residents, which serve to increase the risk for negative outcomes.\textsuperscript{116} This theory suggests that urban places in particular have higher incidences of negative behavior because they are comprised of strangers who do not know one another intimately and because they segregate groups of people into specialized segments that compete rather than cooperate with one another.\textsuperscript{116} Therefore, this theory would predict that adolescent risky sexual behavior, and potential pregnancy, are more common in urban settings than suburban or rural because competition and anonymity among residents in cities make child-rearing and supervision less effective.\textsuperscript{116}
6.2. **Aim Two**: To determine if community factors (race, ethnicity, single parent households, education and poverty) are associated with communities of high adolescent pregnancy prevalence.

Although the urbanism theory is a potential explanation for clusters of communities with high adolescent pregnancy prevalence, there are also reasons to expect that geographic location has more nuanced influences on adolescent pregnancy. It is most natural to look first at some of the known or hypothesized community factors. This study adjusts for five such factors: race, ethnicity, education, poverty, and single parent households.

Findings determined the geographical clustering of communities with high adolescent pregnancy prevalence persisted after adjusting for community factors education, race, ethnicity and poverty. This indicates that these factors individually may not have as great of an impact on adolescent pregnancy prevalence. The additional cluster identified in Syracuse is a subtle shift from the single cluster found in all of the spatial analyses. The larger Syracuse cluster is most likely still present; it just dropped slightly below the two-cluster result in terms of significance. While a different intervention approach could be utilized in this specific area, since only one additional cluster was observed after adjusting for these factors, it could be assumed these factors alone do not make a significant contribution on the prevalence of adolescent pregnancy within New York State communities.

Family structure is an important family context in that children grow up usually having primary relationships with one or two biological parents and with or without older
and younger siblings. Family contextual or structural characteristics have been shown to be related to an adolescents’ risk of pregnancy. With respect to parents’ marital status, research consistently shows that living with a single parent is related to adolescents having sexual intercourse at younger ages. Studies have gone beyond the bivariate relationship to show that single or divorced parents’ more permissive sexual attitudes, lesser parental supervision, and parents’ own dating activity help explain why adolescents in some single parent families are at increased risk of pregnancy. In this study, four clusters emerged after adjusting for the community factor single parent households. This suggests that these communities have a lower percentage of single parent households, and if it were not for this low percentage, they would experience a high prevalence of adolescent pregnancy. This was confirmed in tracking the new cluster in part of the City of Watertown, which identified single parent households as being the least contributing factor.

In 2010, 63 percent of children residing in poor families in New York State lived with a single parent. Between 2006 - 2010, 36 percent of families with children under 18 in both the Western region and state were headed by one parent, which exceeded the national rate of 33 percent. Within the City of Rochester alone, 68 percent or roughly two out of three families were headed by single-parents. Roughly two-thirds of New York City’s poor families with children are families headed by a single mother. Single mother families also constitute two-thirds of New York City’s poor families with children. New York City children living in single mother families endure a 54.0 percent poverty rate. In this study, New York City and part of the City of Rochester were identified as having communities with high adolescent pregnancy prevalence.
Although these clusters were not specifically tracked, based on the above statistics and identification of high prevalence of adolescent pregnancy among these communities, it can be suggested that single parent households may have a large contribution among the factors analyzed in these communities.

Applying the spatial scan statistic as done in this study to adolescent pregnancy clusters allows a better understanding of the extent to which geographic patterns of adolescent pregnancy can be explained by important community factors. Documenting risk for adolescent pregnancy (observed vs. expected) by geographic community after each factor adjustment and tracking changes within the factors provides a useful approach, analogous to the methods used in non-spatial statistics (e.g., Cox regression). While traditional non-spatial analysis provides greater clarity as to the precise contribution of each risk factor, this complementary approach has the advantage of highlighting specific geographic locations.

When adjusting for the six community factors together, all significant clustering of communities with high adolescent pregnancy prevalence disappeared. This suggests that high prevalence of adolescent pregnancy in these clusters was due to a combination of community factors, including race, ethnicity, education, poverty, and single parent households. This study demonstrates that a multitude of community factors are associated with adolescent pregnancy. In addition, focusing prevention efforts on one level of community deprivation, such as education, poverty, race, ethnicity or single parents, may not be sufficient in communities to prevent or reduce adolescent pregnancy prevalence. As no one single community factor is found to be highly associated with
adolescent pregnancy, a comprehensive, community-based approach is recommended to prevent adolescent pregnancy.

Community concern over the high prevalence of adolescent pregnancy has led to a variety of interventions and efforts to lower the prevalence. Until the last decade, however, most of these efforts have been in the form of discrete programs, such as a school-based curriculum in one area of a community, a self-esteem building in another or a health clinic in yet another part of the community. While these programs may be effective in reaching their specific goals, this study indicates that preventing adolescent pregnancy requires multi-component models that involve strategic alliances in the process. As a result, local and state health departments and public health professionals need a more comprehensive strategy to target adolescent pregnancy, which requires broad community involvement and an integrated set of prevention strategies. Such strategies will promote a trend toward greater community collaboration and integration, to assist in solving the problem of adolescent pregnancy within a community.

To build a comprehensive, community-based approach to prevent adolescent pregnancy four key elements are desired: community engagement, comprehensive strategies, sustainability and evaluation. The CDC/ATSDR Committee for Community Engagement developed a working definition of community engagement. Loosely defined, community engagement is the process of working collaboratively with and through groups of people affiliated by geographic proximity, special interest, or similar situations to address issues affecting the well-being of those people. It often involves partnerships and coalitions that help mobilize resources and influence systems, change relationships among partners, and serve as catalysts for changing policies,
programs, and practices. Community engagement is the initial process of strategic alliance formation, involving multiple stakeholders to describe the need, based on information such as that found in this study. With the community engaged, programs can use these resources to discuss potential approaches, develop a shared sense of ownership of the effort and effectively plan a multifaceted intervention. Examples of mobilizing a community include developing relationships with government sectors, school districts and local businesses and hosting communitywide brainstorming sessions.

Comprehensive, community-based programs should be implemented using a combination of interventions to simultaneously address community factors, such as poverty, ethnicity, race, education and single parent households. It is also important to ensure coordination across interventions, to be certain needs are being met and strategies are effective. Embracing diverse norms and values of cultures and racial groups is also critical in implementing a successful community approach. Sustainability of a program may involve instituting structures that oversee efforts across a community. If a community is experiencing success in reducing the risk for adolescent pregnancy, the community must remember not to become complacent. Continual effort will be needed to sustain the reduction in adolescent pregnancy, even after intense interventions are completed. If the sustainability of a community approach is not considered, the problem of adolescent pregnancy may rise again in a new generation of adolescents. Lastly, evaluation is critical to ensure the community wide strategies and interventions are successful, to identify if areas where the program may not be successful, or to identify the emergence of new trends.
Reviews of the literature find shortcomings in studies of teen pregnancy prevention programs. Studies lacked sufficient sample size. Few included long-term follow up. Programs were conducted as demonstration projects, with a maximum of resources and support; very few have been replicated in less ideal circumstances. A number of studies also lacked an experimental design. None of the programs had large, sustained effects on adolescent sexual behavior, contraceptive use, pregnancy, and childbearing rates. Some educational programs reported modest positive effects on delay of sexual initiation and less effect on contraceptive use or pregnancy rates. Among the multi-component programs, some have shown an effect in reducing sexual risk taking, some have not. The most intensive programs were usually the most effective. Effects tended to disappear when programs were stopped.

An example of a successful community-based approach to adolescent pregnancy is that of Minnesota’s Teen Pregnancy Prevention Project. The Teen Pregnancy Prevention Project of Minnesota (TPPPM) developed out of an interagency team into a public-private partnership housed in Minnesota Planning, a state agency. It was governed by a steering committee of representatives of state agencies and local and private agencies. TPPPM was a demonstration project with the mission of reducing teen pregnancies in Minnesota. There were four specific project goals: increasing collaboration at the community and state levels on teen pregnancy prevention issues; increasing effective and replicable teen pregnancy prevention programming; increasing leadership on the state level for the prevention of teen pregnancy and the support of teen parents; and, strengthening policies in Minnesota related to adolescent pregnancy prevention and parenting. To accomplish these goals, four strategies were employed: 1)
implementation of community-based collaborations (Cross Cultural Pregnancy Prevention Peer Education Program, Tri County Coalition on Adolescent Pregnancy Prevention, Youth Health Collaborative and Youth Issues Network); 2) development of statewide professional membership organization, including private, nonprofit, and public service organizations joined to strengthen polices and programming related to adolescent pregnancy prevention and parenting. Services included service provider trainings, annual conference, quarterly newsletter and public policy, public education, and awareness efforts; 3) conducted research and information activities. The program collected and disseminated statistics and objective information on adolescent health and teen pregnancy to increase the efficacy of prevention programming and encourage the exchange of information among local programs, state government and communities; and, 4) provided technical assistance. A variety of training and consultation services were provided to communities and task forces. Evaluation of the TPPPM determined this program achieved a number of its goals. The program increased collaboration on teen pregnancy issues among staff at both the state and local levels. Depth and coherence were added to existing adolescent pregnancy prevention programming, as evidenced in youth development measures and enhanced sexuality communication among youth. By developing community-based comprehensive programs, such as seen in this example, alliances within the community are strengthened to enact change in the social environment, subsequently reducing the risk for adolescent pregnancy.
6.3. **Aim Three:** Determine whether the risk for adolescent pregnancy prevalence changed among communities from 2000 to 2010.

Table 5.6. Model B found education (p<.0001), race (p<.0001), ethnicity (p<.0005), single parent households (P<.0001) and poverty (P<.0001) to be highly significant (p<.001). This finding is consistent with findings from the spatial analysis. The factor foreign born citizens was also determined to be a significant risk for adolescent pregnancy (p<.0095). The generalized linear mixed model showed that, poverty, single parent households, ethnicity, foreign born citizens and race, were all positively associated with adolescent pregnancy. A large number of residents in an adolescent’s community with at least a Bachelor’s degree are found to be a protective factor in this analysis. This factor may serve to buffer adolescents from the potential risks of adolescent pregnancy. These results are in accordance with the spatial model, which identifies the impact of these factors on the risk for adolescent pregnancy among New York State communities. This model underlines the importance of adjusting for confounding community effects when investigating the risk for pregnancy among adolescent females.

In comparing Model A to Model B, the one factor that was not significant at both time intervals, is foreign born citizens. New data from the Census Bureau show that the nation’s foreign-born population reached 40 million in 2010, the highest number in American history. Immigrants come to America for many reasons, such as a desire to join relatives or to access public...
services. As a result, immigration remains high even during a prolonged period of economic weakness. New York State is one of a handful of states with the largest immigrant populations. New York State had more than 800,000 foreign born residents in 2010. New York accounts for roughly 10 percent of the nation’s immigrants.\textsuperscript{125}

New York City has traditionally been the gateway where many immigrants first arrive and settle. Less often recognized, however, is that a substantial number of immigrants, more than 200,000, make their home in the upstate metropolitan areas of Buffalo, Rochester, Syracuse, Albany, Glens Falls, and Utica.\textsuperscript{125} This influx of immigrants upstate has gone relatively unnoticed, masked by the negligible growth in the region’s overall population.\textsuperscript{126} Upstate immigrants have come from a very different set of countries than their downstate counterparts and represent a more varied mix of backgrounds and skills. One out of every five children under 18 in the U.S. is an immigrant or has immigrant parents, and the number of children and youth in immigrant families has expanded almost seven times faster than the number in U.S. born families.\textsuperscript{126}

This influx of the foreign born population includes adults who may lack a high school degree, specialized skills and live in poverty.

Many factors influence migration decisions, such as the desire to be with relatives, the political freedoms in this country, and American public services. These things do not change, even during a prolonged period of economic stagnation. Therefore, immigration is driven in part by social networks of friends and family who provide information about conditions in the United States and often help new immigrants after they arrive. As the immigrant population grows, it creates pressure for more immigration. Economically, times were better at the beginning of the decade, and
immigrants, for the most part, were viewed as assets to the labor force and society. Today, the country, and state are experiencing economic distress, unemployment, and immigrants are more likely to be seen as competitors for jobs and live in poverty.

As this study identified clusters of communities with high adolescent pregnancy prevalence in primarily urban cities, the landscape among cities within the state and nation is worth noting. Residential segregation, or the spatial and social distribution of two or more population groups within a city area, has been hypothesized as a fundamental cause of the racial disparities found in health outcomes. Studies examining the causes of racial segregation mainly focus on three themes: segregation resulting from racial discrimination, from own-race preference and from economic differences. The City of Buffalo was considered one of the highest ranked cities for the most segregated between black and white residents. In New York City, black residents are considered hyper-segregated. Fang and colleagues examined impact of local race concentration on mortality using the U.S. census and mortality data from New York City. They found that within a residentially segregated city, whites and blacks living in ZIP code areas where they were the predominant race had lower mortality rates than persons of other races within their community and persons of their same race who resided in other areas. Hispanics, the second most segregated group, are highly segregated in a number of cities, including New York City. These findings are consistent with what has come to be termed as ethnic density effect and may be explained by health enhancing consequences of residing in residential areas where opportunity exists for greater social support and protection from direct prejudice. While directly targeting racial segregation as part of adolescent pregnancy prevention efforts may be
beyond the scope of an intervention, being sensitive to this dynamic may provide greater utility to a prevention problem.

Overall, findings suggest adolescents residing in certain New York State communities are at an increased risk for adolescent pregnancy. The increased risk is in part due to a combination of community factors, such as the percentage of residents living in poverty, the number of black and Hispanic residents residing in the community, the number of single parent households, and a lower percentage of community residents with a Bachelor’s degree or higher. From 2000 to 2010, the number of communities that observed an increase in risk for adolescent pregnancy, based on identified community influences, has gotten larger. Figure 5.9 depicts the communities that observed an increase in the level of risk for adolescent pregnancy over the last decade. This map demonstrates the trend of risk among New York State communities over the last decade as getting worse. Pockets of communities at greater risk were identified across the state, with the largest number of communities observed in the Adirondack/North Country, Western Region, Catskills and Central Region. The map also suggests that the risk of adolescent pregnancy has the expectation of moving from a primarily urban issue to a concern for poorer rural areas.

Findings demonstrate that while certain clusters of communities are currently experiencing high adolescent pregnancy prevalence, there are many more communities at risk for high prevalence of adolescent pregnancy. While targeting clusters of communities with high adolescent pregnancy prevalence is a priority, those communities at high risk, that may not be exhibiting a high prevalence, cannot be disregarded. If these communities are not treated and provided with community-based pregnancy prevention
programs, they could present a problem in the future. All of these communities demonstrating risk are in need of comprehensive, coherent, multi-faceted adolescent pregnancy prevention programs, to prevent adolescent pregnancy from becoming a bigger problem in their community, in the future.

6.4. Future Recommendations

Geographic deprivation of New York State communities, as well as risk and protective factors of these communities should be considered when designing and targeting adolescent pregnancy prevention programs, as discussed above. To have a greater effect on prevention, programs must address community factors. An awareness of the influence of factors such as race, ethnicity, poverty, single parent households, education and foreign born citizens, will provide public health professionals with a better understanding of their influence on communities. To potentially have a greater impact on pregnancy prevention, programs must address antecedents related to social disorganization as well. While more research and evaluation is certainly warranted, programs that recognize meaningful strategies including community-wide, comprehensive, sustainable intervention strategies will be the most effective. As adolescent pregnancy is a complex problem, with multiple underlying factors, strategic alliances or partnerships among multiple sectors of a community are most likely to have the most significant impact on reducing or preventing adolescent pregnancy.
6.5. Limitations

This research relied primarily on two data sources, the U.S. Census and New York State Vital Statistics birth data. The U.S. Census does not provide a total enumeration of a population; therefore, results may only provide a cross-sectional view of the population.\textsuperscript{133} The same is true for vital statistics data sets, which may be undercounted. Additionally, the 2010 ACS estimates at the ZIP code level used in this study, were not those officially released by the U.S. Census, as this data was not available at the time of this study. However, based on the conversion calculations of the ACS 2010 Census developed as part of this study using the tested and approved relationship file to coordinate geographic location, these estimates should mirror those released by the U.S. Census in the near future.

By using ZIP codes as a community reference, there may be instances in New York State where communities at risk may be hidden. For example, there is only one ZIP code for Troy, New York, which encompasses a large geographic area with diverse characteristics and populations. Within this ZIP code, North Troy is found to have a similar demographic profile as that of the other urban cities identified as a risk for adolescent pregnancy; therefore, it would be expected that this community would be identified as a cluster for high adolescent pregnancy prevalence. However, because only one ZIP code is assigned for all of Troy, which represents both North Troy and surrounding areas, we cannot tease out the heterogeneity of this community and it is subsequently defined as not posing a risk for adolescent pregnancy.

It is important to note that the geographic boundaries of the clusters detected in this study are approximations of the true clusters. Meaning, while the general location of
the cluster is known, the exact boundaries are uncertain. As with any ecological analysis, it cannot be said that the whole adolescent population residing within a cluster of communities with high adolescent pregnancy prevalence is at the same risk for pregnancy. Adolescents have varying levels of risk, which may depend on their individual characteristics and behaviors; however, the presence of the clusters of communities suggest added community factors exist in that area and pose an additional risk.

Despite these limitations, this approach provides an objective and practical method to identify community factors of risk for the high prevalence of adolescent pregnancy. This research extends an understanding of the geographic distribution of adolescent pregnancy in communities and may aid in the understanding of associations related to adolescent pregnancy.

6.6. Conclusion

In summary, this study identified clusters of communities with high adolescent pregnancy prevalence by geographic location throughout New York State. The clusters of communities with high adolescent pregnancy prevalence persisted after adjustments for race, education, ethnicity, poverty and single parent households, individually. This suggests that these factors, alone, are not a main contributing community factor for risk of adolescent pregnancy. Conversely, when these factors were combined, all geographic clustering was removed. These findings suggest that, in communities where adjustment eliminated the geographic clusters and reduced the risk of adolescent pregnancy, factors
such as race, ethnicity, education, poverty and single parent households, may be contributing causes of the community disparities.

The specific clusters of communities with high adolescent pregnancy prevalence were identified in urban cities, specifically, New York City, Buffalo, Rochester, Syracuse, Niagara Falls, Utica, Binghamton, Albany and Schenectady. In many parts of the state, more people now live in suburbs than in cities. The middle-class exodus over the past few decades has led to a general worsening of conditions in cities, such as greater levels of poverty, higher levels of vacant housing, a greater percentage of female headed households with children, and a higher percentage of adults with less than a high school diploma when compared to surrounding suburban neighborhoods. Given these dynamics, statistically significant clusters of communities with high adolescent pregnancy prevalence could be suggested to result from negative effects of urban features and larger, denser and more heterogeneous cities, which increase the potential for negative outcomes, such as adolescent pregnancy. Given the fiscal climate today, and in the foreseeable future, the large, dense, and diverse urban cities will continue to foster social disorganization and alienation among residents, which serve to increase the risk for adolescent pregnancy. The problem of adolescent pregnancy also has the expectation of moving from a primarily urban city issue to a concern of poorer rural areas.

The characteristics of the community setting in which an adolescent resides provide a context that has the potential to influence adolescent pregnancy and to inform the development and expansion of accessible and effective community-based prevention and treatment services. The use of the methods presented in this study, demonstrate how researchers and public health professionals can monitor adolescent pregnancy, evaluate
the effectiveness of statewide or locally based interventions and generate hypotheses about the underlying causes of communities on the risk of adolescent pregnancy. Providing timely and relevant information to public health professionals on how and why adolescent pregnancy varies geographically is necessary to ensure interventions are targeting the right communities and using the most effective and efficient approach.
REFERENCES


