The Effects of Maternal Age on the Prevalence of Autism

Melissa A. Bilbao[1], Alexander D. Castro[2], Tyler A. Rigazio[3]

[1] Department of Mathematics and Statistics, California State Polytechnic University

Pomona, CA 91768, USA.

[2] Department of Mathematics, University of Illinois at Urbana-Champaign

Urbana, IL 61801, USA.

[3] Department of Mathematics and Statistics, The University of Maine

Orono, ME 04469, USA.

August 2008

Abstract

Autism's cause is unknown, but suggested causes are often attributed to genetic or environmental factors. This research examines whether advancing maternal age contributes to the increasing prevalence of autism. To test our hypothesis, we create a deterministic model. Values are generated representing the proportion of offspring expected to be diagnosed with autism, provided their mother belongs to a specific age class. The results show that women of the age class consisting of ages 40-44 are more susceptible, by nearly 20 percent, to having a child who will be diagnosed with autism. Using our model, projections are made about the prevalence of autism in future populations, specifically, for the United States and California. These projections predict a continued increase in the prevalence of autism. A linear regression model is also used to statistically confirm that maternal age is affecting the increase in the prevalence of autism.

A Introduction

A.1 Characteristics and Symptoms of Autism

Autism belongs to a class of behavioral disorders known as autism spectrum disorders (ASDs). ASDs affect people from all backgrounds, regardless of race, ethnicity, or socioeconomic status, but appear four times more often in boys than in girls [?, ?]. Children diagnosed with an ASD typically learn, focus, and react in ways contrary to unaffected individuals. The symptoms of autism (characterized as high or low functioning) typically manifest themselves prior to 18 months, providing for earliest diagnosis at the age of two [?]. While the disorder lasts the duration of a person's life, autistic individuals vary greatly in their capacity for interaction and learning.

Deficiencies in learning and interaction lead to the enrollment of many autistic individuals in special education classes. Autistic people tend to learn in a systematic manner; repetition is crucial to maintaining a relaxed environment conducive to processing information [?, ?]. This systemizing quality has been suggested as an explanation for why many autistic individuals possess innate mathematical talent. According to psychologist Simon Baron-Cohen, autistic people are very logical and good at processing things governed by rules [?]. They are good at dealing with order and if this order becomes disrupted, behavioral outbursts are triggered. Autistic individuals generally exhibit more left-brained characteristics than right [?]. Because of this, individuals are deficient in traits governed by the right brain, such as the capacity for sympathy, emotional connection with others, interaction, and expression. Autism can severely hinder an individual's ability to progress within society. Their abilities however, can be astounding, as supported by the existence of autistic savants [?].

A.2 Treatment

As of yet, no absolute cause for autism has been determined. Despite the lack of certainty regarding cause, treatments are available to lessen the severity of symptoms and medication can suppress the occurrence of outbursts. Behavioral and educational aids exist for therapists to help autistic patients improve social and language skills in a structured learning environment. Occupational therapy benefits autistic individuals by improving the quality of their social life; it provides an environment fit for developing fine motor skills and the capacities for sympathy, coping, and self-help. Sensory Integration Therapy is used to facilitate the development of the nervous system's ability to process sensory input in a "normal" way. Many forms of autism-specific

therapies exists, but most have improved relationship-building skills and vocalization of needs and desires as their primary goal [?].

Aside from medication and therapy, life-style changes have become an increasingly popular method for treating autism. The Gluten-Free, Casein-Free Diet is based upon the hypothesis that proteins in foods containing gluten and casein are absorbed by an autistic individual's body in such a way that they act as false opiate-like chemicals. Eliminating these foods seems to regulate the behaviors of autistic individuals [?].

A.3 A Review of Current Research

Autism research is presently focused on a psychological, neurological, and genetic aspect. Most research is not focused on cause, but rather regulation. However, this focus has not limited research analyzing autism's cause(s) and many mathematical models have been created to improve the understanding of this complex disorder [?].

Some research related to our own has been done. A study in Northern California, consisting of a small sample size of singleton children at a local hospital, used proportional hazard regression to determine the risk associated with increasing age in having an autistic child [?]. The data supported that greater maternal and paternal age increased the risk of having a child with autism. To be more specific, an Israeli study focused only on paternal age [?]. The categorical measure used allowed for the modeling of the effects of paternal age on autism through a system of nonlinear equations. A sensitivity analysis was used and the relationship between paternal age and autism was determined to be monotonically increasing. Like the Northern California study, Israel used a small sample size and were unable to find an impact due to maternal age. Instead, they concluded that increased paternal age was more likely to contribute to autism-prevalent mutations.

Aside from those that consider genetic factors and analyze risk ratios, other mathematical models have tried to find autism-specific triggers. For example, a paper by Janusonis detailed a regulation model analyzing the role of serotonin blood levels in autistic individuals [?]. It is speculated that there is a link between autism and statistically higher 5-hydroxytyptamine levels. A mathematical model was constructed to see if failure of molecular mechanisms of 5-HT could account for autistic abnormalities. However, studies failed to find a link between autism and serotonin-levels within the brain of autistic individuals. Glutathione has also been studied [?]. The model examined the properties of glutathione metabolism and it's regulations of enzymes by simulating a metabolic profile of autism under increased oxidative stress levels.

Research done in California used multivariate models to calculate risk estimates in a sample of 4381 children [?]. It was found that a higher autism risk existed for boys, the children of African American women, and multiple births. Mothers of higher education and greater age were shown to be more susceptible to having a child with the disorder. The study also concluded that children born to immigrant mothers (of California) had a lower-risk for being diagnosed with autism than others born in California. Connections between environmental factors and the disorder were made.

Another recent study included statistical analyses of spontaneous mutations that occur in autistic males of families who are considered low-risk (high risk families consist of the offspring of unaffected females who carry a new causative mutation and transmit the mutation to their offspring in a dominant manner) [?]. This was analyzed using a binomial distribution and a log-likelihood function.

Other genetic mutations, such as clustering of alleles, might account for autism's increased prevalence [?]. Of late, the controversy regarding the mandatory vaccinations and their effects, if any on autism's increase, has taken center stage. In the past, vaccines contained mercury [?]. While the mercury has been removed from the vaccines, examination of autistic individuals' baby teeth have shown higher mercury content than that measured in individuals who have not been diagnosed with autism.

A.4 Motivation for Study

The lack of knowledge regarding autism may suggest that it is a relatively modern disorder, however, historical descriptions of symptoms similar to those used to diagnose autism date back to the 18th century. Autism was identified as a developmental disorder in 1943 by psychiatrist Dr. Leo Kranner [?]. His criteria for diagnosis included: lack of affective contact, desire for sameness, fascination with objects, and mutism or non-communicative language before 30 months of age. Since Kranner's description, the criteria for diagnosis have been modified many times [?].

The age and criteria for diagnosis are ever-changing and only partially account for the increased incidence of autism within society. Autism is becoming significantly more prevalent with time, especially within the United States, as seen in Figure 9. Within 10 years, autism prevalence has risen from $\frac{1}{10,000}$ to $\frac{1}{150}$ [?]. Another explanation and the topic of this research attributes much of the increase to women's growing tendencies to

postpone pregnancy. Many social factors accounting for this trend include: the choice to pursue higher education, establish a career, and marry later in life [?]. Unfortunately, with delayed maternal age comes susceptibility to genetic mutations during fetal development and other complications. The mother's body becomes more fragile and she becomes more likely to require a cesarean section [?]. Because maternal age appears to have such a significant effect on child birth, the impact of maternal age on the prevalence of autism diagnoses will be investigated. Our model is used to estimate the proportion of children who will be diagnosed with autism, given their mother's age class, in the United States and California.

Compared to prior models and their intentions, our model assumes a different role. This model neither focuses on a small sample of the population nor does it analyze risk ratios for having a child who will be diagnosed with autism; it makes predictions at the population level. It seeks to make a conjecture about which maternal ages are more susceptible to having a child who will be diagnosed with autism by using a nonlinear model.

We organized this paper as follows. In Section 2, we discuss the data we collected and analyze its reliability. Section 3 gives an overview of the deterministic model we create to analyze the impact of maternal age on the increasing prevalence of autism in both California and the United States. The results of this model are detailed in Section 4, as is a check of the model's reliability, and projections for the growth of the autistic population in the United States and California for years 2010 to 2050. An alternative approach in which we construct a linear regression model confirms the findings of our deterministic model; this approach is explained in Section 5. People influential in the success of this project are acknowledged in Section 6. Following the references, the appendix includes the current DSM-IV diagnostic procedure for autism and all relevant figures.

B The Data

Our research is driven by available data. Sources report the number of autistic individuals by age (between the ages of 6 and 22) who are enrolled in special education programs within public schools per year [?, ?]. The maximum number of autistic individuals within the public school system is determined by birth year after some reorganization of the data provided. For example, if it is recorded that there are 25 autistic children of age 6 in 1992 (meaning they were born in 1986), but 30 individuals of age 7 in 1993, it is concluded that there are actually 30 autistic children born in 1986. The trend in the autism data provided is for the value to increase (as more individuals of the same birth year are diagnosed) and then decrease (as individuals leave the public school system). Because this decline is not observed in some of the data, it is realized that after 1990, the maximum number of autistic individuals of a given birth year cannot be determined absolutely; the maximum value available is therefore used.

The data set used is underestimated and affected by underreporting [?]. Autistic individuals attending private and home schools and those outside of the special education program are not included. Only those who meet the eligibility criteria for autistic disabilities are accounted for. This is problematic because the working definition of autism varies from state to state, as do diagnostic procedures and the services available to those with autism. Despite the limitations of this data, they are the most comprehensive and reliable data available. From them, autism's prevalence by age, state and birth year can be estimated.

Autism was first included in the Diagnostic and Statistical Manual for Mental Disorders in 1980 [?, ?]. Because of its infancy as a recognized disorder, many misdiagnoses have occurred and continue to occur. Also, public schools were not required by law to report the number of autistic individuals until 1992; thus, these data are further underestimated. By examining the United States as a whole, despite the lack of uniformity in diagnosis from state to state, it is believed that excellent diagnostic procedures and states with better record keeping will offset the more flawed systems and a nationwide average can be analyzed.

For California and the United States, census data are used to calculate the total number of offspring expected to attend public school based upon their year of birth and the age class to which their mother belongs [?, ?]. Information is gathered about the size of the female population for specific age classes and their respective birth rates [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?]. By multiplying these together the expected number of offspring is calculated. Multiplying this result by the proportion of children who are enrolled in public school for year $i + 6$ (this year approximates when offspring born in year i would begin entering the public school system), the total number of offspring expected to attend public school is found.

From the data, it is seen that the autistic population and the reproductive female population both increase exponentially from 1983 to 1993. Based upon a curve of best fit for the data, the number of autistic individuals in the United States is increasing at a rate of approximately 0.197 each year, as seen in Figure 9. The national female population is growing at a rate of 0.008 per year.

Figure 9: Number of Autistic Individuals Born in the United States in 1983 to 1993

C The Model: A Deterministic Approach

It is observed that the number of children being diagnosed with autism within the United States is increasing exponentially [?]. For the purpose of this project, all possible causes of autism, not linked to maternal age, are attributed to the environment. This provides for near isolation of maternal age so that its precise role in the increasing prevalence of autism can be examined. The nonlinear model we create to test the hypothesis that advancing maternal age is contributing to the increase in the number of individuals being diagnosed with autism per year is seen in Equation 23:

$$Y_i = \sum_{j=1}^{6} x_{i,j}\beta_{i,j}\tau_i\gamma_j e^{\alpha i} \tag{23}$$

If we let $x_{i,j}$ represent the number of females in year i, belonging to age class j; $\beta_{i,j}$ be the birth rate in year i, unique to women of age class j; τ_i be the proportion of offspring born in year i who are expected to attend public school; γ_j be the age class-specific proportion of offspring who will be diagnosed with autism; and α be the contribution of environmental factors to the increase in autism, then Y_i will represent the number of children born in year i who will be diagnosed with autism, for $i = 1, ..., 11$. Year 1983 is represented by $i = 1$ and year 1993 is represented by $i = 11$. Exempting γ_j and α, our model's parameter values are obtained from data collected by various organizations. The years spanned by each age class is shown in Table C. It is assumed that only women of these age classes can reproduce. We also assume that the likelihood of a mother sending her child to public school is independent of whether or not the child is autistic. This is also independent of which age class she belongs to. This means that τ_i remains constant for a given year i.

Table 3: Age Classes

Age Class j	1	2	3	4	5	6
Age Range	15-19	20-24	25-29	30-34	35-39	40-44

Table 4: Definitions

Y_i	The number of children born in year i who will be diagnosed with autism
$x_{i,j}$	The number of women during the ith year in the jth age class
$\beta_{i,j}$	The birth rate specific to women during year i of age class j
τ_i	The proportion of offspring born in year i who will attend public school
γ_j	The proportion of offspring born to mothers of age class j who are expected to be diagnosed with autism
α	The environmental factor to the increase in autism

By including an exponential component in our model to account for the increase in autism due to environmental factors (especially misdiagnosis), we eliminate any natural exponential growth in the autistic population due to overall growth in the number of females reproducing each year. Because of this, when solving for α, we adjust the resultant value by the annual growth of the female population. For example, if α is 0.119 and 0.01 is the approximate growth of the female population per year, then 0.01 is subtracted from 0.119 so that the autistic population is growing as would be expected, given that the female population is increasing exponentially. The final α value that would be used to calculate the γ values is therefore 0.109.

The summation over j for each i value generates a system of 11 equations of 7 unknowns. To be biologically feasible, certain restrictions exist on the values that α and γ_j can assume. The proportion of offspring born to women of age class j who will be diagnosed with autism, γ_j, must lie between 0 and 1. In the case of the United States (1983-1993), the environmental contribution to autism's increase, α, must lie between 0.008 and 0.197, where 0.008 is the growth of the female population and 0.197 is the observed growth in the autistic population.

For California, α is contained within the interval (0.021, 0.220). This variation in the values α can assume makes our model nonlinear. However, for ease of analysis, we force our model to be linear by fixing α. To determine the best value at which we fix α, we use MATLAB to test α values in the ranges given, incrementing the value by 0.001. Each α value produces values for γ_j and generates an exponential curve that estimates the growth in autism. With this in mind, the best α value and the one that relays the correct information about the extent that maternal age is impacting the increase in autism, is the one that generates the curve that best reflects the observed growth in autism. MATLAB is used to find each α value, display resultant γ_j values calculated using the method of nonnegative least squares, and evaluate the error associated with each choice. The method of nonnegative least squares is used to guarantee that the values for γ_j are nonnegative and therefore provide for biological interpretation. The error is calculated according to the method of mean square error, which is one of many methods that could be used to calculate the deviation of the model's values from the observed. Error must be minimized to achieve accuracy in applying our model. Once the α with the smallest error is found, the growth in the female population is subtracted from its value. We then take the γ_j values that correspond to this new α value and we are able to see how women are affected by their age in their likelihood of having a child who will be diagnosed with autism.

D Results and Analysis

D.1 California: A Rather Ideal Locale for Autism Study

More so than most states, California takes great interest in autism [?, ?, ?]. Because of this, autism data is well maintained through the Department of Developmental Services and much autism research is being done within the state. In an epidemiology report prepared by the University of California, Davis confirmed that from 1983 to 1993, the increase in autism was fairly independent of misclassification into other developmental disorder categories (i.e. mental retardation), younger diagnosis age, and influx of residents seeking better autistic care services [?]. The number of autistic individuals is shown in Figure 10. Because diagnosis standards were consistent, it is expected that autism in California has a lower overall environmental component than can be expected for the United States [?].

Figure 10: The Number of Autistic Individuals Born in California in 1983 to 1993

Table 5: γ Values for California

Age Class	1	2	3	4	5	6
γ	0	0	0	0	0	0.2071

From MATLAB we calculate α to be 0.090 as seen in Figure 11. The resultant γ values are shown in Table 5. For women of ages 40-44, the proportion of their offspring expected to be diagnosed with autism is 0.2071, provided maternal age is the only contributing factor. However, because maternal age is not the only contributing factor, these numbers simply show that women of ages 40-44 are more likely to have an autistic child than if they are susceptible to environmental factors alone. The other age classes appear unaffected.

Figure 11: Mean Square Error vs. α for California

In comparing the Mean Square Error of 83.8639 to the data, it is seen that the error is rather insignificant as the number of autistic individuals for years 1983-1993 range from 280 to 2248. Thus, the γ values are reasonable because they provide results that do not stray too far from the observed data.

D.2 United States: 1983-1993

We perform a similar study for the United States. It is expected that the α value should be larger than California's because within the United States diagnosis is changing more rapidly than in California. Since diagnosis is nonuniform across the United States and the exponential component adjusts for uniform diagnosis, a higher Mean Square Error is expected [?]. For the United States α is found to be 0.114, as seen in Figure 12. The respective γ values are shown in Table 6.

Figure 12: Mean Square Error vs. α for the United States

Table 6: γ Table Values for the United States

Age Class	1	2	3	4	5	6
γ	0	0	0	0	0.0364	0.1218

Since the Mean Square Error is 540.1421 and the range of data is 2573 to 16759, the error is insignificant. Similarly to the California data, this verifies that the model is reasonable since it also does not stray too much from the observed data. The error in the United States however, is much larger than California's error; this explains why there is now a value associated with the age class consisting of ages 35-39. Despite this, a trend still exists within the United States supporting the notion that with advancing maternal age comes increased susceptibility to having an autistic child.

D.3 Comparison

The γ values can now be used to test the reliability of the model by predicting the number of autistic children within the United States and California from 1994 to 2000 [?]. United States census data allow for the calculation of the number of offspring and matrix multiplication predicts the number of autistic individuals expected to enter the special education program within the public school system. Predictions will be made using two methods, one in which it is assumed that environmental factors remain constant after 1992 ($\alpha = 0$) and one in which the effects of environmental factors increase exponentially.

D.3.1 California

Using γ values from the California table and assuming α is zero, the predicted number of autistic individuals for years 1994-2000 and the deviation from the actual data is given in Table 7. Also, the equation used for the predictions follows:

$$\sum_{j=1}^{6} x_{i,j}\beta_{i,j}\tau_i\gamma_j e^{\alpha i} \tag{24}$$

for $i = 1...7$, where $i = 1$ denotes year 1994 and $i = 7$ denotes year 2000.

Table 7: Observed/Predicted Autistic Population for California in the Absence of Environmental Factors ($\alpha = 0$).

Year	Observation	Prediction
1994	2501	2259.2
1995	2592	2374.7
1996	2758	2523.4
1997	3023	2595.7
1998	3200	2689.7
1999	3328	2719.3
2000	3465	2975

Figure 13: Predicted Autistic Population for California, 1994 to 2000 ($\alpha = 0$)

The Mean Square Error is 449.89, and as seen in Figure 13, the number of autistic births per year increases from 2501 to 3465 over this time period; hence the error is reasonable. The predicted values consistently underestimate the actual number of autistic individuals within California. Since the Mean Square Error is reasonable, the expected values are relatively close to the original data.

D.3.2 The United States

Using γ values for the United States and assuming that α is zero, the predicted number of autistic individuals for years 1994-2000 and the observed population values are seen in Table 8 and in Figure 14.

Table 8: Observed/Predicted Autistic Population Values for the United States in the Absence of Environmental Factors ($\alpha = 0$)

Year	Observation	Prediction
1994	17744	19300
1995	19715	20140
1996	20034	21150
1997	21425	22020
1998	22163	23000
1999	21292	23980
2000	21315	29950

Figure 14: Predicted Autistic Population for the United States, 1994-2000 ($\alpha = 0$)

The Mean Square Error for this data set is 3799.2 and seen in Figure 14 the number of autistic individuals from 1994 to 2000 ranges from 17744 to 21315. This shows that despite relatively large population values, the error is too great to conclude that the model is accurate without incorporating increasing environmental factors. A contributing factor to this error is the absence of a true peak in the data for the number of autistic individuals; it is quite likely that many autistic individuals are unaccounted for due to under-reporting.

D.3.3 Varying α

When running the California simulations in MATLAB, the best α value is 0.029, as seen in Figure 15. The corresponding Mean Square Error is 127.1652. The predicted values are shown in Table 9. From this, we can conjecture that environmental factors are still having an effect on the prevalence of autism.

For the United States, the best α value is found to be zero. The analysis has been detailed above.

D.4 Projections

Now that the model shows some semblance of reliability, projections for the years 2010-2050 can be made. Using census population projections for California, the projected birth rates and the projected rates of public school enrollment using a best fit line, the following changes in the prevalence of autism are projected and shown in Figure 16. Projections for the United States are made using the same process as was used for California. This is illustrated in Figure 17.

Over the next 50 years, it is projected that the autistic population will continue to grow in both California and the United States. California will have a net increase between 2010 and 2050 of 5,106 autistic individuals, while the United States autistic population will increase by nearly 192,050 individuals.

Figure 15: Mean Square Error vs. α for California

Table 9: Prediction Table for the California with $\alpha = .029$ with Mean Square Error of 127.1652

Year	Observation	Prediction
1994	2501	2325.7
1995	2592	2516.5
1996	2758	2752.7
1997	3023	2915.0
1998	3200	3109.4
1999	3328	3236.1
2000	3465	3644.5

D.5 The Model: A Linear Regression Approach

In this section we will perform hypothesis tests to discern the affects of maternal age on the prevalence of autism is by taking a linear regression approach. This approach has the potential to confirm the results of the deterministic model. Note that by using linear regression, the values given for γ_j may not be biology feasible (in that they are not contained within the interval [0,1]). The possibility also exists that there will be no noticeable trend in the γ_j values. In fact, if from our data we can show that any error is randomly and normally distributed, then the resultant values for γ_j should be similar to the results obtained if we were to use the method of regular least squares. The linear regression model is shown in Equation 25.

$$Y_i = \rho_0 + \sum_{j=1}^{6} x_{i,j}\rho_j + \epsilon_i \tag{25}$$

If we let $x_{i,j}$ represent the number of offspring born in year i to mothers of age class j who are expected to attend public school in year $i+6$, ρ_j be the effect of maternal age of mother in age class j, and ϵ_i be the error in estimating the number of autistic children each year, then Y_i, is the number of children born in year i who will be diagnosed with autism. The response (or dependent) variable is Y_i and the predicted variables (or independent) are $x_1, x_2, \cdots, x_6$. These parameters are organized in Table 8.

Table 10: Definitions

Y_i	The number of children born in year i who will be diagnosed with autism
$x_{i,j}$	The number of offspring in year i to mothers of age class j who will attend public school
ϵ_i	The difference between the predicted and observed values

The first thing to be tested with our linear regression approach is if there is a correlation between maternal age and the number of offspring diagnosed with autism. Our null hypothesis is: $H_0 : \rho_1 = \rho_2 = \cdots = \rho_6 = 0$. If we are able to reject this hypothesis (show that at least one ρ_i is not equal to zero), then we will have statistically proven that maternal age is indeed playing a role in the prevalence of autism. Using MINITAB we obtain the regression equations for the United States (Equation 26) and California (Equation 27) and their respective P-values.

The regression equation for the United States is

$$Y = -26328 - .0177x_1 + .028x_2 - .0099x_3 + .0047x_4 + .0548x_5 + .238x_6. \tag{26}$$

The regression equation for California is

$$Y = 711.5 + .0063x_1 + .00669x_2 - .0156x_3 - .0307x_4 + .0976x_5 + .084x_6. \tag{27}$$

Figure 16: Projected Autistic Population from 2010-2050 Born in California

Figure 17: Projected Autistic Population from 2010-2050 Born in the United States

The P-values obtained for both California and the United States are 0.000. Since these P-values are small, we reject the null hypothesis. In addition to this, we can conclude that each age classes has a varying role as seen in examining the coefficients of x_j in the regression equations. For example, x_1 is playing a negative role in the prevalence of autism in the United States. This means that for every unit by which x_1 is increased, the population of autistic individuals will decrease by the coefficient of x_1. Note that MINITAB also generates graphs of the residuals, as shown in Figure 18 and in Figure 19, *(figures for California are located in the Appendix)*. We are able to conclude that the error is independently and identically distributed, meaning that the obtained ρ_j values would have also resulted if we had used the method of regular least squares. It is observed that the ρ_j coefficients do not follow a trend. Despite this, they do show that each age class plays a relevant role in contributing to the prevalence of autism. Because of this, we can conclude that our deterministic model is valid. We do not utilize linear regression or use regular least squares to solve for our γ_j values in the deterministic model because by using non-negative least squares, we are able to obtain solutions that are biologically feasible.

Another hypothesis further confirms that maternal age is playing a role in the increasing prevalence of autism. In this case, we see if the ρ_j values are equivalent. The null hypothesis is: $H_0 : \rho_1 = \rho_2 = \cdots = \rho_6$. If we can reject this hypothesis (show that at least two ρ values are not equal), then this further emphasizes that maternal age is affecting the number of offspring who will be diagnosed with autism. The reduced model for this hypothesis is shown in Equation 28.

$$Y_i = \rho_0 + \sum_{j=1}^{6} x_{i,j}\rho_j + \epsilon_i,\ C\rho = H \tag{28}$$

where,

$$C = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \rho = [\rho_1 \cdots \rho_6]', \text{ and } H = [0 \cdots 0]'.$$

Because we know the response and predicted variables, the C matrix, β matrix, and the H matrix, we can apply the F-test which is seen in Equation 29.

$$F = \frac{\frac{SSE(R) - SSE(F)}{s}}{\frac{SSE(F)}{n-p}} \tag{29}$$

Since the proper matrices were configured from our response and predicted variables we are able to use MATLAB to do the proper matrix multiplication. The resultant F- and P-values are shown in Table 11.

Table 11: F- and P-values for the United States and California obtained from MINITAB

	F-value	P-value
United States	247.32	0.00001
California	269.588	0.00001

The resultant P-values obtained from the F-test are small; thus, we can reject our null-hypothesis and conclude that there exists at least two unequal ρ_j values. From this, we can more strongly assert that maternal age is in fact contributing to the increase in the prevalence of autism.

Figure 18: Residuals vs. Fit Graph for the United States

Figure 19: Normal Probability Plot for the United States

E Conclusion and Discussion

We seek to determine whether or not advancing maternal age is playing a role in the increasing prevalence of autism. To do this we construct a nonlinear deterministic model from which we are able to estimate parameter values corresponding to the proportion of offspring who are expected to be diagnosed with autism, given that their mother belongs to a specific age class [?, ?, ?]. Analyzing California and the United States, the model is kept at the population level. Maternal age is separated from environmental factors, providing for a more focused analysis. It is found that as mother's age, their susceptibility to having a child who would be diagnosed with autism increases dramatically. The results show a recurring trend that support the idea that, in the absence of environmental factors, increased maternal age increases a woman's susceptibility to having an autistic child. In California, a state that is fairly advanced and consistent in its diagnosis of autism, women of age class 40-44 have a 20.71 percent chance of having an autistic child [?]. For the United States, values are found for both the 35-39 year age class (3.64 percent) and the 40-44 year age class (12.18 percent). Despite the appearance of a value in the age class 35-39, it is believed that the 40-44 year age class is the only one of significance and that the value for the 35-39 year age class can be attributed to error in the data set. Because of this, it can also be concluded that maternal age has little or no perceivable impact on the likelihood of having an autistic child for women of ages 15-39. After checking the reliability of our model, we are able to make projections about the prevalence of autism in future populations. It is projected that autism will continue to increase over the time spanning 2010-2050.

We construct a linear regression model as another way to determine whether or not maternal age impacts the likelihood of having a child who will be diagnosed with autism. Two hypotheses are tested and the results obtained confirm the results of our deterministic model; that is, maternal age does indeed play a role in the increasing prevalence of autism. It is important to note that as much as we tried to isolate maternal age by attributing all other causes to the environment, we cannot be certain that we actually succeeded in doing so. There are many other factors that could be contributing to the values associated with each age class; for instance, if it is paternal age rather than maternal age affecting the number of offspring being diagnosed with autism, our model would report similar findings. Therefore, because autism is likely caused by multiple factors other than maternal age, the results of this study allow for the suggestion that maternal age should be more closely studied so as to determine its precise role in the increasing prevalence of autism.

As evidenced by this project, much remains to be learned about autism. To expand upon this research, it would be interesting to see if women in the age class 45-49 have a greater susceptibility than those of age class 40-44, if the impact of maternal age varies by nation or race or even if paternal age has a greater impact than maternal age [?]. Also, a sibling study that examines the increased risk associated with having an older brother or sister who is autistic would be a beneficial and intriguing avenue to pursue. To pursue this research however, would require data collection methods to be much improved over their current state, especially with regard to the autistic population. If data is collected on the mother's age when she gave birth to a child diagnosed with autism, better statistical studies could be done to determine if maternal age is playing the role suggested by this research. Also, a standardized diagnosis system needs to be implemented. Much work remains to be done in this field and it is hoped that this paper inspires many others.

F Acknowledgement

We would like to thank Baojun Song for all the help, support and motivation to continue to work on the project even when progress seemed impossible. Your professionalism forced us to keep our antics in check, but your humor provided the Autismers with unwavering spirit. Thank you for you willingness to engage in fisticuffs (look it up) so that we could retain some semblance of sanity and pursue our original idea. You are a ridiculous man with amazing math skill. Also, many thanks to Xiaohong Wang, allowed us to explore

the realm of autism and mathematics as we wished. Your explanations, model ideas, wonderful laugh, high spirits, "WHAAAAAAAT?!", and your willingness to participate in MTBI gossip made our experiences more enjoyable. The addition of Kimberly Rude for those short, yet oh so awesome, 2 weeks of non-stop data finding, model arguing, and amazing-greek-food-eating united Team Autismers. You always knew the right questions to ask. Thank you Naala Brewer for lifting our spirits when we were down and for proof-reading the 2938479208374 drafts. Thank you Steve Tennenbaum for motivating the project and keeping our wild imaginations in check. Thank you Dan Hollenback for giving us the data that made this project possible. And finally, thank you to Carlos Castillo-Chavez for starting this program and being more than meets the eye. Your interest and motivation in our project kept us helped us persevere. The Autismers love you ALL! We would further like to acknowledge, the MTBI/SUMS Summer Undergraduate Research Program which is supported by The National Science Foundation (DMS-0502349), The National Security Agency (DOD-H982300710096), The Sloan Foundation, and Arizona State University.

References

[1] Centers for Disease Control and Prevention. Achievements in Public Health, 1990-1999: Impact of Vaccines Universally Recommended for Children – United States, 1990-1998. *MMWR Morb Mortal Wkly Rep*, 48:241–248, 1999.

[2] Centers for Disease Control and Prevention. Ten Great Public Health Achievements - United States, 1900-1999. *MMWR Morb Mortal Wkly Rep*, 48:241–243, 1999.

[3] SL Plotkin and SA Plotkin. *Vaccines*, chapter 1, pages 1–12. W.B. Saunders Company, 3 edition, 1999.

[4] AB Coffield, MV Maciosek, JM McGinnis, JR Harris, MB Caldwell, SM Teutsch, D Atkins, JH Richland, and A Haddix. Priorities among recommended clinical preventive services. *American Journal of Preventive Medicine*, 21(1):1–9, 2001.

[5] WA Orenstein, RG Douglas, LE Rodewald, and AR Hinman. Immunizations in the united states: Success, structure, and stress. *Health Affairs*, 24(3):599–610, 2005.

[6] P Davies. Antivaccination web sites. *JAMA*, 288(14):1717–1718, Oct 2002.

[7] SP Calandrillo. Vanishing vaccinations: Why are so many americans opting out of vaccinating their children? *University of Michigan Journal of Law Reform*, 37:353–440, 2004.

[8] RT Chen. *Vaccines*, chapter 49, pages 1144–1163. WB Saunders Company, 3 edition, 1999.

[9] Centers for Disease Control and Prevention. Some Common Misconceptions about vaccination and how to respond to them. `http://www.cdc.gov/vaccines/vac-gen/6mishome.htm`, Accessed July 12, 2008.

[10] YA Maldonado. Current controversies in vaccination: Vaccine safety. *JAMA*, 288(24):3155–3158, 2002.

[11] A Wakefield, S Murch, A Anthony, J Linnell, D Casson, M Malik, M Berelowitz, A Dhillon, M Thomson, and P Harvey. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, 351(9103):637–641, 1998.

[12] US Food and Drug Administration. Thimerosal in vaccines. `http://www.fda.gov/CbER/vaccine/thimerosal.htm`, Accessed July 8, 2008.

[13] E Gangarosa, A Galazka, C Wolfe, L Phillips, R Gangarosa, E Miller, and R Chen. Impact of anti-vaccine movements on pertussis control: the untold story. *The Lancet*, 351(9099):356–361, 1998.

[14] R Jacobson, P Targonski, and G Poland. A taxonomy of reasoning flaws in the anti-vaccine movement. *Vaccine*, 25(16):3146–3152, 2007.

[15] GA Poland and RM Jacobson. Understanding those who do not understand: a brief review of the anti-vaccine movement. *Vaccine*, 19:2440–2445, 2001.

[16] RM Wolfe. Content and design attributes of antivaccination web sites. *JAMA*, 287(24):3245–3248, 2002.

[17] RD Silverman. No more kidding around: Restructuring non-medical childhood immunization exemptions to ensure public health protection. *Annals of Health Law*, 12:277–294, 2003.

[18] J Thompson, S Tyson, P Cardhigginson, R Jacobs, J Wheeler, P Simpson, J Bost, K Ryan, and D Salmon. Impact of addition of philosophical exemptions on childhood immunization rates. *American Journal of Preventive Medicine*, 32(3):194–201, 2007.

[19] DA Salmon, M Haber, EJ Gangarosa, L Phillips, NJ Smith, and RT Chen. Health consequences of religious and philosophical exemptions from immunization laws: Individual and societal risk of measles. *JAMA*, 282(1):47–53, 1999.

[20] DA Salmon and AW Siegel. Religious and philosophical exemptions from vaccination requirements and lessons learned from conscientious objectors from conscription. *Public Health Reports*, 116(4):289–296, 2001.

[21] SB Omer, WKY Pan, NA Halsey, S Stokley, LH Moulton, AM Navar, M Pierce, and DA Salmon. Nonmedical exemptions to school immunization requirements: Secular trends and association of state policies with pertussis incidence. *JAMA*, 296(14):1757–1763, 2006.

[22] JG Hodge. School vaccination requirements: Legal and social perspectives. *NCSL State Legislative Rep*, 27:1–14, 2002.

[23] D Khalili and A Caplan. Off the grid: Vaccinations among homeschooled children. *The Journal of Law, Medicine & Ethics*, 35(3):1073–1105, Jul 2007.

[24] T May and RD Silverman. 'Clustering of exemptions' as a collective action threat to herd immunity. *Vaccine*, 21(11-12):1048–1051, 2003.

[25] A d'Onofrio, P Manfredi, and E Salinelli. Vaccinating behaviour, information, and the dynamics of SIR vaccine preventable diseases. *Theoretical Population Biology*, 71(3):301–317, 2007.

[26] A Maayan-Metzger, P Kedemfriedrich, and J Kuint. To vaccinate or not to vaccinate - that is the question: why are some mothers opposed to giving their infants hepatitis B vaccine? *Vaccine*, 23(16):1941–1948, 2005.

[27] A Wroe, A Bhan, P Salkovskis, and H Bedford. Feeling bad about immunising our children. *Vaccine*, 23(12):1428–1433, 2005.

[28] Thomas C. Schelling. *Micromotives and Macrobehavior*. W. W. Norton and Company, 1978.

[29] WO Kermack and AG McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, 115(772):700–721, 1927.

[30] JJ Gart. The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics*, 24(3):557–566, 1968.

[31] F Ball. Deterministic and stochastic epidemics with several kinds of susceptibles. *Advances in Applied Probability*, 17(1):1–22, 1985.

[32] H Lacayo and NA Langberg. The exact and symptotic formulas for the state probabilities in simple epidemics with m kinds of susceptibles. *Journal of Applied Probability*, 19(1):1–9, 1982.

[33] CM Kribs-Zaleta and JX Valesco-Hernández. A simple vaccination model with multiple endemic states. *Mathematical Biosciences*, 164:183–201, 2000.

[34] S Rushton and AJ Mautner. The deterministic model of a simple epidemic for more than one community. *Biometrika*, 42(1-2):126–136, 1955.

[35] RK Watson. On an epidemic in a stratified population. *Journal of Applied Probability*, 9(3):659–666, 1972.

[36] NG Becker and K Dietz. The effect of household distribution on transmission and control of highly infectious diseases. *Mathematical Biosciences*, 127(2):207–219, 1995.

[37] F Ball, D Mollison, and G Scalia-Tomba. Epidemics with two levels of mixing. *Annals of Applied Probability*, 7(1):46–89, 1997.

[38] F Ball and P Neal. A general model for stochastic SIR epidemics with two levels of mixing. *Mathematical Biosciences*, 180:73–102, 2002.

[39] FG Ball, T Britton, and OD Lyne. Stochastic multitype epidemics in a community of households: Estimation of threshold parameter r and secure vaccination coverage. *Biometrika*, 91(2):345–362, 2004.

[40] RM Anderson RM May. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences*, 72(1):83–111, 1984.

[41] DJ Murrell, U Dieckmann, and R Law. On moment closures for population dynamics in continuous space. *Journal of Theoretical Biology*, 229:421–432, 2004.

[42] D Hiebeler. Moment equations and dynamics of a household SIS epidemiological model. *Bulliten of Mathematical Biology*, 68(6):1315–1333, 2006.

[43] NG Becker and DN Starczak. Optimal vaccination strategies for a community of households. *Mathematical Biosciences*, 139(2):117–132, 1997.

[44] DH Zanette and M Kuperman. Effects of immunization in small-world epidemics. *Physica A: Statistical Mechanics and its Applications*, 309:445–452, 2002.

[45] Z Lu, X Chi, and L Chen. The effect of constant and pulse vaccination on SIR epidemic model with horizontal and vertical transmission. *Mathematical and Computer Modelling*, 36(9):1039–1057, 2002.

[46] PA Briss, LE Rodewald, AR Hinman, AM Shefer, RA Strikas, RR Bernier, VG Carande-Kulis, HR Yusuf, SM Ndiaye, and SM Williams. Reviews of evidence regarding interventions to improve vaccination coverage in children, adolescents, and adults. *American Journal of Preventive Medicine*, 18(1S):97–140, 2000.

[47] PA Gross, AW Hermogenes, HS Sacks, J Lau, and RA Levandowski. The Efficacy of Influenza Vaccine in Elderly Persons: A Meta-Analysis and Review of the Literature. *Annals of Internal Medicine*, 123(7):518–527, 1995.

[48] MJ Fine, MA Smith, CA Carson, F Meffe, SS Sankey, LA Weissfeld, AS Detsky, and WN Kapoor. Efficacy of pneumococcal vaccination in adults. A meta-analysis of randomized controlled trials. *Archives of Internal Medicine*, 154(23):2666–2677, 1994.

[49] J Ward, J Cherry, S Chang, S Partridge, and H Lee. Efficacy of an acellular pertussis vaccine among adolescents and adults. *New England Journal of Medicine*, 353(15):1555–1563, 2005.

[50] S Black, H Shinefield, B Fireman, E Lewis, P Ray, JR Hansen, L Elvin, KM Ensor, J Hackell, and G Siber. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. *The Pediatric Infectious Disease Journal*, 19(3):187, 2000.

[51] K O'Brien, L Moulton, R Reid, and R Weatherholtz. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *The Lancet*, 362:355–361, 2003.

[52] N Becker. Estimation for discrete time branching processes with application to epidemics. *Biometrics*, 33(3):515–522, 1977.

[53] D Hiebeler. Populations on fragmented landscapes with spatially structed heterogeneities: Landscape generation and local dispersal. *Ecology*, 81(6):1629–1641, 2000.

[54] DE Hiebeler and AK Criner. Partially mixed household epidemiological model with clustered resistant individuals. *Physical Review E*, 75(2), 2007.

[55] O Diekmann, J Heesterbeek, and J Metz. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990.

[56] K Dietz. The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research*, 2(1):23–41, 1993.

[57] J Li JM Hyman. An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations. *Mathematical Biosciences*, 167:65–86, 2000.

[58] C Castillo-Chavez, JX Velasco-Hernandez, and S Fridman. Modeling contact structures in biology. *Frontiers of Theoretical Biology, Lecture Notes in Biomathematics*, 100:454–491, 1994.

[59] Ying-Hen Hsieh, P Driessche, and Lin Wang. Impact of travel between patches for spatial spread of disease. *Bull. Math. Biol.*, 69(4):1355–1375, May 2007.

[60] R Cohen, S Havlin, and D ben Avraham. Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24):247901:1–4, 2003.

[61] M Murray and Z Rasmussen. Measles Outbreak in a Northern Pakistani Village: Epidemiology and Vaccine Effectiveness. *American Journal of Epidemiology*, 151(8):811–819, 2000.

[62] S van den Hof, CMA Meffre, MAE Conyn van Spaendonck, F Woonink, HE de Melker, and RS van Binnendijk. Measles Outbreak in a Community with Very Low Vaccine Coverage, the Netherlands. *Emerging Infectious Diseases*, 7(3 Supplement), 2000.

[63] J Puvimanasinghe, C Arambepola, and N Abeysinghe. Measles Outbreak in Sri Lanka, 1999-2000. *The Journal of Infectious Diseases*, 2003.

[64] C Stein-Zamir, N Abramson, H Shoob, and G Zentner. An outbreak of measles in an ultra-orthodox Jewish community in Jerusalem, Israel, 2007 - An in-depth report. *Eurosurveillance*, 13:1–4, 2008.

[65] JL Richard, V Masserey-Spicher, S Santibanez, and A Mankertz. Measles outbreak in Switzerland - An update relevant for the European football championship (EURO 2008). *Eurosurveillance*, 13(8), 2008.