Who's Afraid of the Big Bad Wolf? Do pedophiles live closer to schools?

Angelo Mele
Department of Economics
University of Illinois, Urbana-Champaign
amele2@uiuc.edu
http://netfiles.uiuc.edu/amele2/www

January 27, 2008

Abstract

I analyze the geographic distribution of sex offenders in Urbana and Champaign and test if they live closer to schools than generic residents. If the spatial distribution of sex offenders is the same as that of population as a whole, we should not observe a sex offender systematically closer to schools than a random resident. Using a simple statistical model I test if, conditioning on the distance from schools, the spatial distributions of sex offenders and residents are the same.

The results show that this is not the case: *on average* sex offenders are *less* likely to locate close to schools than other people. However, I show evidence that there is some heterogeneity among *individual schools*, some being sex offenders "magnets".

1 Introduction

On May 12, 2006 the channel CBS2 located in Los Angeles, aired a report by David Goldstein, in which the reporter showed several previously convicted sex offenders living very close to schools.[1] Given the high level of recidivism for this type of crime one can speculate that this location is not random, but those sex offenders have deliberately chosen an apartment close to a school for predatory purposes. There are laws prohibiting sex offenders to live within a certain distance from a school. For example, in Illinois:[2] *It is unlawful for a child sex offender to knowingly reside within 500 feet of a school building or the real property comprising any school that persons under the age of 18 attend.* Given these regulations, one may wonder how this can happen and if this behavior is common.

In this work I analyze the spatial distribution of sex offenders in the cities of Urbana and Champaign, Illinois, to detect if a representative sex offender lives closer to schools than the representative resident. I perform my test using data from the Illinois Sex Offender Registry. The Illinois Compiled Statutes (730 ILCS 152/115 (a) and (b)) mandate that the Illinois State Police ("ISP") establish and maintain a statewide Sex Offender Database, accessible on the Internet, identifying persons who have been convicted of certain sex offenses and/or crimes against children and must register as a Sex Offender.[3] The Registry is available online at `http://www.isp.state.il.us/sor/`.

I have information about the current address of 124 sex offenders.[4] Figure 1 shows the area of Urbana-Champaign with the location of all sex offenders (red dots) and schools (the blue houses). The proximity of a sex offender to a school does not seem rare: for example in Urbana there are two sex offenders living less than half a mile away from the Yankee Ridge Elementary School, located at 2102 S Anderson Street.

[insert Figure 1 here]

The test is based on a class of models, the point source models, used in spatial epidemiology to detect if the cases of a certain disease show spatial association with several sources of environmental pollution (see Lawson (2001), Diggle and Rowlingson (1994)). For example the model is used to test if cancer cases are clustered around a nuclear plant (the source). In my application the model will test if the location of sex offenders is systematically clustered around the schools.

The intuition of my test is simple. The fact that we observe several sex offenders living close to schools is not conclusive, since individuals choose where to live according to

[1] The video is availble online at `http://www.cbs2.com/video/?id=18427@kcbs.dayport.com`

[2] Availble at `http://www.isp.state.il.us/docs/720ilcs51193.pdf`

[3] Persons required to register as Sex Offenders are persons who have been charged of an offense listed in Illinois Compiled Statutes 730 ILCS 150/2(B) when such charge results in one of the following:

(a) A conviction for the commission of the offense or attempt to commit the offense,

(b) A finding of not guilty by reason of insanity of committing the offense or attempting to commit the offense, or

(c) A finding not resulting in an acquittal at a hearing for the alleged commission or attempted commission of the offense.

More details are contained in the Appendix or in the website.

[4] There are 177 registered offenders, but some of them are currently in jail, homelesses or noncompliant with the registration duties. We had to exclude those from the analysis.

neighborhood's amenities and socio-economic composition: the presence of a school in the neighborhood increases the appeal of a certain location. If sex offenders behave like the representative resident, ceteris paribus they will choose to live close to a school. On the other hand, if the spatial distribution of sex offenders reflects the same preferences of the rest of individuals, on average we should not observe sex offenders closer to schools than the rest of population.

The model postulates that the location of residents and sex offenders follow two independent spatial Poisson point processes. A Point Process is a statistical model describing how points locate in a two-dimensional space. The intensity function of the process is the pointwise expected number of points per area and the intensity measure is the expected number of points per area. The process is a Poisson point process with intensity $\lambda(x)$ if: 1) for any sub-area of the city, the total number of points n of the process is a draw from a Poisson distribution with mean equal to the intensity measure; 2) conditional on the draw n, the locations are identically and independently distributed over the area with density proportional to the intensity function.

The locations of sex offenders are a realization of a process with intensity $\lambda(x) = \rho\lambda_o(x) f(x - x^0; \theta)$, where ρ is a parameter indicating the ratio of sex offenders and non sex offenders, $\lambda_o(x)$ is the intensity of generic residents and $f(x - x^0; \theta)$ is the variation in the intensity reflecting the distance from the schools locations x^0.

If there is no clustering of the sex offenders around schools the function $f(x - x^0; \theta)$ should be equal to 1: I specify a parametric model for f and I test if the parameters θ's guarantee the latter condition.

The ideal dataset for this test contains the location of sex offenders, schools and residents (used as controls): unfortunately it is very difficult to access residential address data, because of confidentiality issues. Therefore I rely on simulation methods. I estimate the intensity function of the population from the Census 2000, Summary Files 1, at the block level, using nonparametric methods. I then simulate 1000 realizations of the estimated process with 1000 locations each, and use those as controls in the point source model. I estimate the parameters (ρ, θ) using maximum likelihood techniques for each sample and obtain an empirical distribution of the parameter estimates that I use for inference.

I use the empirical distribution to compute the mean of the parameters, the standard errors and perform tests of significance.

The baseline results show than on average the probability that a sex offender lives very close to a school is *lower* than the same probability for a generic individual, the former probability being close to zero. Furthermore the former probability is an *increasing* function of the proximity to a school. Similar results are obtained when considering each type of school (elementary, middle and high) separately. This may be due to the effectiveness of the law preventing sex offenders from living close to schools or it may be just the results of individual preferences for location, but I cannot discriminate between the two hypothesis using this test.

Since the baseline results refer to the average clustering around all the schools taken together, I also estimate the individual schools' contributions to the average. It appears that in general high schools are more appealing for sex offenders and middle schools are less, and that the average results are driven mainly by a few individual schools. I conclude that there is some heterogeneity: some schools act like magnets while other schools are safe.

The general lesson is that while on average there is no evidence of sex offenders living closer to schools than generic residents in Urbana and Champaign, I cannot exclude that some schools are more exposed to sex offenders recidivism.

2 A Simple Statistical Model

A spatial point process is a statistical model describing how points locate in a two-dimensional space. I model the metropolitan area as a set $\mathcal{S} \subseteq \mathbb{R}^2$ and the agents' locations as the realization of a spatial point process X defined in $\mathcal{S}$. A realization $\{x_1, ..., x_n\}$ of the process X is a set of locations in the metropolitan area. Alternatively a spatial point process can be characterized as a random variable $N(A)$ which measures the number of points of the process in any area $A \subseteq \mathcal{S}$. The process is driven by the *intensity function*

$$\lambda(x) = \lim_{|dx| \to 0} \left\{ \frac{\mathbb{E}N(dx)}{|dx|} \right\} \tag{2.1}$$

where dx is the infinitesimal ball around x and $|dx|$ is its area. The intensity function is the expected number of points of the process X at location x. It is a pointwise analog of the concept of geographic density (in expectations). In order to recover the expected number of points in a set $A \subseteq \mathcal{S}$ one integrates the intensity function over A

$$\mathbb{E}N(A) = \Lambda(A) = \int_A \lambda(x)\, dx \tag{2.2}$$

The integral above is called *intensity measure.* The process is *simple* if there are no coincident points and *inhomogeneous* if the intensity function is non-costant over $\mathcal{S}$.

The model postulates that the locations of sex offenders and the rest of population follow two independent Poisson point processes.

A point process X defined on $\mathcal{S}$ is a *Poisson Point Process* with intensity $\lambda(x)$ if

1. for any $A \subseteq \mathcal{S}$, $N(A) \sim Poisson(\Lambda(A))$

2. conditional on $N(A) = n$, the events are identically and independently distributed over A according to the density $h(x) = \lambda(x)/\Lambda(A)$

The first condition is that the number of points in any subset of the metropolitan area is a draw from a Poisson distribution with mean the intensity measure of the process $\Lambda(A)$. The second condition assumes that there is no interaction among locations.

The framework I use for my test is a point source model borrowed from the spatial epidemiology literature (see Lawson (2001), chapter 7). This class of models is often used to detect if the locations of cases of a specific disease are clustered around a source of environmental pollution. For example, Diggle and Rowlingson (1994) model the spatial distribution of cancer cases in South Lancashire and test if a disused industrial incinerator has an impact on the high level of cancers in the region. The same model is used to test if asthma cases recorded in North Derbyshire were clustered around three industrial plants. In my application I will treat the locations of sex offenders as cases and the rest of the population as controls.

Suppose to have a dataset containing the location of sex offenders x_i, $i = 1,..,n$ and the location of non sex offenders x_i, $i = n+1,...,n+m$ in a metropolitan area. I assume that the locations of generic individuals are a realization of a simple Inhomogeneous Poisson point process with spatially varying intensity $\lambda_o(x)$, while the locations of sex offenders are a realization of a simple inhomogeneous Poisson point process with intensity

$$\lambda(x) = \rho\lambda_o(x) f\left(x - x^0; \theta\right) \tag{2.3}$$

In this formulation ρ is a parameter indicating the ratio of sex offenders and non sex offenders,[5] $\lambda_o(x)$ is the intensity of generic individuals and $f(x - x^0; \theta)$ is the variation in the intensity reflecting the distance from the schools locations x^0. Suppose there are q schools, then the function $f(x - x^0; \theta)$ is modelled as follows

$$f\left(x - x^0; \theta\right) = \prod_{k=1}^{q} g\left(x - x_k^0; \theta\right) \tag{2.4}$$

Therefore the total change in intensity of sex offenders $f(x - x^0; \theta)$ is a multiplicative function of the change arising from each school $g(x - x_k^0; \theta)$.

In order to estimate the model (2.3) - (2.4) notice that, conditional on the realization of the $n+m$ locations, the individuals can be any of two types: sex offender or not. So, conditioning on the $n+m$ locations, the variable indicating if an individual is a sex offender behaves like a Bernoulli variable and therefore the probability that the individual living at location x is a sex offender is

$$p(x; \rho, \theta) = \frac{\rho f(x - x^0; \theta)}{1 + \rho f(x - x^0; \theta)} \tag{2.5}$$

Using (2.5), the log-likelihood function in the usual Bernoulli form is

$$L(\rho, \theta) = \sum_{i=1}^{n} \ln p(x_i; \rho, \theta) + \sum_{i=n+1}^{n+m} \ln\left[1 - p(x_i; \rho, \theta)\right] \tag{2.6}$$

The parameters estimates are the maximizers of (2.6).

3 Empirical Strategy

3.1 Functional Forms and Estimation

In the baseline specification the function $g(x - x_k^0; \theta)$ is modeled as a negative exponential function of the distance of the point from the school

$$g(d_k; \theta) = 1 + \theta_1 \exp\left[-\theta_2 d_k\right] \tag{3.1}$$

[5] As long as the researcher can choose the controls, this parameter has no interest and just reflects this choice.

where $d_k = \left[(x - x_k^0)'(x - x_k^0)\right]^{-1/2}$. The coefficient θ_1 represents the scale/shift factor while θ_2 represents the decay factor driving the intensity. If sex offenders are systematically clustered around schools we expect that both θ_1 and θ_2 are positive: sex offenders are more likely to live close to a school than the generic resident and the probability that a sex offender lives at a specific location is a decreasing function of the distance from the school. On the other hand, if the law is effective in keeping the sex offenders away from schools, we expect the opposite sign for θ_1. When there is no difference between the spatial distribution of sex offenders and the residential distribution we expect both the coefficients to be zero, or at least $\theta_1 = 0$

This model is very parsimonious in terms of parameterization, since it requires to estimate two parameters to identify the function $f(x - x^0; \theta)$ and one parameter ρ in order to compute the probabilities. I also show results based on an alternative specification with the same number of parameters

$$g(d_k; \theta) = 1 + \theta_1 \exp\left[\theta_2 \ln(d_k) - d_k\right] \tag{3.2}$$

which allows for a peaked distance decay probability of location.

An implicit assumption in (3.1) is that the vector of parameters (θ_1, θ_2) is the same for each school. An alternative model would have a different parameter for each individual elevation function, i.e. $g(d_k; \theta_k)$, so the number of parameters needed to identify $f(x - x^0; \theta)$ would be $2q$. [6] I present several results based on such a model

$$g(d_k; \theta_k) = 1 + \theta_{1,k} \exp\left[-\theta_{2,k} d_k\right] \tag{3.3}$$

Given the small sample of sex offenders I cannot estimate the full model, which would require 53 parameters (I have 26 schools), but I show separate estimates for each type of school.

Under the parameterization in (3.1), the log-likelihood function can then be written as

$$\begin{aligned} L(\rho, \theta) &= n \ln \rho + \sum_{i=1}^{n} \sum_{k=1}^{q} \ln\left\{1 + \theta_1 \exp\left[-\theta_2 d_k\right]\right\} \\ &\quad - \sum_{i=1}^{n+m} \ln\left[1 + \rho \prod_{k=1}^{q} \left\{1 + \theta_1 \exp\left[-\theta_2 d_k\right]\right\}\right] \end{aligned} \tag{3.4}$$

Notice that under H_0 we have $(\theta_1, \theta_2) = (0.0)$ and $f(x - x^0; \theta) = 1$, therefore the log-likelihood in (3.4) becomes

$$L_0(\rho) = L(\rho, 0) = n \ln \rho - (n + m) \ln(1 + \rho) \tag{3.5}$$

with maximum likelihood estimate $\widehat{\rho}_0 = n/m$.

In order to test if sex offenders are systematically closer to schools than the generic inhabitant of Urbana-Champaign, I perform a simple likelihood ratio test. If there is an increase or decrease in the intensity of sex offenders around the schools, then the parameters (θ_1, θ_2) will be nonzero. My test's null hypothesis is thus $H_0 : \theta = 0$ with test statistic

$$T = 2\left[L\left(\widehat{\rho}, \widehat{\theta}\right) - L_0\left(\widehat{\rho}_0\right)\right] \tag{3.6}$$

[6] Another easy extension would consider spatial covariates in the g functions. This may require the estimation of too many parameters though.

that I compare with the critical values of the appropriate Chi-Square distribution.[7] The maximization of the log-likelihood function is performed using a simplex method with the Nelder-Mead algorithm.

3.2 Data

At the end of March 2007, I collected data from the Illinois Web Registry of Sex Offenders for the cities of Urbana and Champaign. The data contain the address of each registered sex offender, some minimal demographic information (gender, race) and the description of the offense. There are 177 compliant sex offenders, but I exclude those currently convicted and the homelesses: this gives a total of 124 sex offenders for the analysis.

The data from schools are collected from the websites of the school districts of Champaign and Urbana,[8] where I found the addresses of each school. There are 26 schools: 19 elementary schools or kindergarten, 4 middle schools and 3 high schools.

In order to estimate the model one would ideally have access to individual location data from the Census or some survey. These data are very difficult to obtain, given confidentiality issues. Therefore I rely on simulation techniques in order to create the location data for non-sex offenders..

3.3 Sample Simulations and Inference

My strategy consists of several steps:

1. I estimate the intensity function $\lambda_o(x)$ from Census data at the block level, using nonparametric techniques
2. I simulate 1000 samples of 1000 points from the process with intensity $\widehat{\lambda}_o(x)$
3. I estimate the parameters and perform the LR test for each sample, obtaining a set of estimates $\left\{\widehat{\theta}_r, \widehat{\rho}_r, T_r, pval_r\right\}_{r=1}^{1000}$
4. I make inference using the single tests and the distribution of the test statistic

In order to recover the intensity $\lambda_o(x)$, I use block-level data from the Census 2000 Summary File 1 (SF1). These data contain the location of the block centroid and the total population of the block. The metropolitan area $\mathcal{S}$ is partitioned in K disjoint blocks and the Poisson assumption implies that the random variables $N(\mathcal{S}_k)$ and $N(\mathcal{S}_l)$ over disjoint regions $\mathcal{S}_k$ and $\mathcal{S}_l$ are independent. It follows $\mathbb{E}N(\mathcal{S}_k) = \int_{\mathcal{S}_k} \lambda_o(x)\, dx$, for any k and the number of points can be rewritten as

$$N(\mathcal{S}_k) = \int_{\mathcal{S}_k} \lambda_o(x)\, dx + u_k$$

[7] Diggle and Rowlingson (1994) provide more details about the test. They perform several simulations in order to check that it is appropriate to use the critical values of the Chi-Square in this setting, concluding positively.

[8] http://www.champaignschools.org/ and http://www.usd116.org/home/

where u_k is a mean zero error, uncorrelated across blocks. There exists a $\overline{x} \in \mathcal{S}_k$ such that $\int_{\mathcal{S}_k} \lambda_o(x)\, dx = \lambda_o(\overline{x}) |\mathcal{S}_k|$ and therefore $N(\mathcal{S}_k) = \lambda_o(\overline{x}) |\mathcal{S}_k| + u_k$.

Assume that $\lambda_o(x)$ is a smooth function and the area of the block $|\mathcal{S}_k|$ is small: for $x \in \mathcal{S}_k$ the following approximation holds

$$N(\mathcal{S}_k) \approx \lambda_o(x) |\mathcal{S}_k| + u_k$$

This allows us to use a kernel regression approach to estimate the expected number of points in $\mathcal{S}_k$,

$$\mathbb{E}[N(\mathcal{S}_k)| x] \approx \lambda_o(x) |\mathcal{S}_k|$$

and thus the function $\lambda_o(x) |\mathcal{S}_k|$ can be estimated as

$$\widehat{\lambda_o}(x) |\mathcal{S}_k| = \sum_{k=1}^{K} \frac{\mathcal{K}_h(x - x_k)}{\sum_{j=1}^{K} \mathcal{K}_h(x - x_j)} n_k$$

where x_k's are the centroids of the census blocks and $\mathcal{K}_h$ is a kernel function. The area $|\mathcal{S}_k|$ is known (from the SF1 data) then I have an estimate of $\widehat{\lambda_o}(x)$, from which I get $\widehat{h}(x)$ to simulate the process and create the samples.

Once I create the samples, I plug them in the maximization routine and find the corresponding parameters estimate by MLE. I perform the LR test for each sample.

I obtain a vector of parameter estimates and tests, one for each sample. I compute the mean of the parameters and take that as my Monte Carlo estimate.

4 Results

In Figure 2 I show the projection of the residential locations in the plane: the longitude and latitude are converted in northings and eastings using a conic projection, and the distance is measured in kilometers. The red dots represent the sex offenders, the blue squares are the schools and the red dots are the centroids of the Census blocks. The position of the blocks gives a rough idea of where the denser areas are.

[insert Figure 2 here]

The estimated intensity surface is shown in Figure 3. The kernel bandwidth is chosen using an MSE minimization procedure described in Diggle (2003): the chosen bandwidth is $h = .428$ km. Notice that areas with higher intensity do not necessarily correspond to the areas in Figure 2 with more block centroids, since the blocks have different population sizes and building structures.

[insert Figure 3 here]

To give a flavor of what is the outcome of my samples simulation procedure I show in Figure 4 two realizations of the process with the estimated intensity. The black dots in the figure represent the simulated locations of the population that I use as controls for my test.

4.1 Baseline Results

The empirical distributions of the parameters are reported in Figure 5 for inspection. The density is estimated using a kernel density estimator

[insert Figure 5 here]

The densities of ρ (a) and θ_2 (c) are reasonably unimodal and close to a normal, but the density of θ_1 shows some irregularity. It may be that 1000 samples are not enough to get a similarly normal distribution.

[insert Table 1 here]

The main results are contained in Table 1. Panel A considers the specification (3.1): I report the mean coefficient, the standard deviation of the empirical distribution and the t statistic. The coefficients are all significant and the sign of θ_1 is negative and close to -1, showing that sex offenders do not cluster around schools. The decay parameter θ_2 is positive, indicating that the probability that a sex offender lives at a certain location increases with the distance of the location from a school.

[insert Figure 6 here]

The result is better explained with a picture: in Figure 6(a) I show the g function implied by the estimated coefficients in Table 1. On average the spatial distribution of sex offenders and rest of the population are the same if the locations we considers are 10 kilometers away from a school. In the immediate proximity of a school the probability of residence for a sex offender is almost null.

The LR test is implemented for each of the samples: I report the rate of rejection for the null hypothesis of no difference between the spatial distributions of sex offenders and other residents. For this specification I reject H_0 in 995 samples. I conclude that the estimated coefficients show a significant difference in the two spatial distribution, the sex offenders being less clustered around schools than the generic residents. The Panel B of Table 1 reports the results based on the specification (3.2) that allows for peaked distance decay. The coefficients are very similar to the previous estimates, but the implied g function is different, as shown in Figure 6(b). These estimates tell us a partially different story: if the location is very close to a school the probability of residence for a sex offenders is almost the same as the other residents. As we move away from the school the probabilty drops quickly to half and around a distance of 2 kilometers starts increasing again. In this second specification I reject the null in all the samples.

The conclusion arising from the two specifications are not very different, in the sense that they both show that the probability of finding a sex offender living close to a school is lower than that of a generic resident. Unfortunately, with this simple test is not possible to discriminate among alternative explanations: is this happening because sex offenders are compliant to the law of it is just because they don't have a strong preference for living close to schools?

4.2 Alternative Specifications

In the following I will use the baseline specification. In Table 2 I present the results by type of school. The coefficients are not always significant and show some interesting patterns. The results in Panel A shows that sex offenders locations do not show any spatial association with the elementary schools since both the shift and the decay parameter are not significant. We reject the null of no change in intensity for sex offenders in 93.2% of the samples.

[insert Table 2 here]

For Middle schools the decay parameter θ_1 is significant and close to -1, showing that sex offenders do not locate in the proximity of middle schools. The percentage of rejections for the LR test falls to 78%. In the last panel the decay parameter is now significant while the shifter is not: this means that very close to high schools, sex offenders have the same spatial distribution of the rest of the population, but the probability of finding a sex offender at a specific location decreases very fast as one moves away from the high school location. The rate of rejection is now only 7%.

[insert Table 3 here]

In tables 3 to 5 I repeat the exercise using the more flexible specification, allowing a different shifter and decay parameter for each school. The main concernt with these specifications is that the number of parameters to estimate increases and I have not enough observations to guarantee precise estimates. One could simply increase the number of observation in the control samples, but if the ratio cases/controls is too low, the model will treat the sex offenders locations as outliers, and they will have very few explanatory power in the likelihood function: therefore increasing the controls without increasing the cases will not solve this problem.

Table 3 shows the estimates for the 19 elementary schoools. The coefficient $\theta_{i,s}$ refers to the parameter $i = 1, 2$ for school $s = 1, ..., 19$. Very few coefficients are significant. Most of the times the decay parameter is positive and significant but the shifter is not. I interpret this outcome as an indication that the probability of finding a sex offender close to any school is not higher than that of a generic resident, confirming the above results. For schools $8, 10$ and 12 the shifting parameter is positive and significant: this means that those schools are sex offenders "magnets". This is not in contradiction with the more aggregated results, since those measure an average effect of the distance from schools, while the more flexible specifications allow the researcher to find each school individual contribution to this average effect. I perform for each school an F test for the joint significance of both the parameters: the test always rejects the null of both parameters being zero.

[insert Table 4 here]

Table 4 provides similar suggestions: the distribution of sex offenders is not more clustered around middle schools than the spatial distribution of residents. For one school the decay parameter is significant and very high in magnitude. The F tests always reject the coefficients are both zero for each school.

[insert Table 5 here]

The high schools 1 and 2 seem to attract more sex offenders than any other school. This means that the average effect in Table 2 is driven mainly by school number 3. Again the F tests are always rejected.

5 Conclusion

I analyze the spatial distribution of sex offenders in Urbana and Champaign, IL to test if the spatial distribution of sex offenders is clustered around schools. I perform a simple test using a point source model: the test consists of measuring if the probability of finding a sex offender around a school is on average higher than the probability of finding a generic resident.

The baseline results show than on average the probability of a sex offender living close to a school is lower than that of the rest of the population. Moreover this probability increases as we move away from the proximity of a schools. Similar results are obtained when considering each type of school (elementary, middle and high) separately.

The baseline results refer to the average clustering around all the schools taken together, but using a more flexible specification I am able to estimate individual schools' contribution to the test. It appears that in general high schools attract more sex offenders and middle schools less and that the average results are driven mainly by a few schools.

The general lesson is that there is no evidence of sex offenders living closer to schools than generic residents in Urbana and Champaign: on average the location probability of a sex offender is lower than that of the average inhabitant the closer we are to a school.

References

[1] Baddeley, Adrian and Turner, Rolf (2005), "spatstat: An R Package for Analyzing Spatial Point Patterns", Journal of Statistical Software, 12(6):1-42

[2] Diggle, Peter (1985), "A Kernel Method for Smoothing Point Process Data", Applied Statistics, 34(2):138-147

[3] Diggle, Peter (2003), "Statistical Analysis of Spatial Point Patterns", Academic Press, London, Second Edition

[4] Diggle, Peter, Zheng, Pingping and Durr, Peter (2005), "Nonparametric estimation of spatial segregation in a multivariate point process: bovine tubercolosis in Cornwall, UK", Applied Statistics, 54(3):645-658

[5] Diggle, Peter and Rowlingson Barry (1994), "A Conditional Approach to Point Process Modelling of Elevated Risk", Journal of the Royal Statistical Society, Series A 157(3):433-440

[6] Kelsall, Julia E. and Diggle, Peter J. (1998), "Spatial Variation in Risk od Diseases: A Nonparametric Binary Regression Approach", Applied Statistics, 47(4):559-573

[7] Lawson, Andrew B. (2001), Statistical Methods in Spatial Epidemiology, Wiley Series in Probability and Statistics, Wiley and Sons

[8] Rowlingson, B.S. and Diggle P.J. (1993), "Splancs: Spatial Point Patterns Analysis Code in S-Plus", Computers and Geosciences, 19:627-655

A Computations

The conditional (on $n+m$) probability that a pedophile lives in location x is computed as follows

$$\begin{aligned} p(x;\rho,\theta) &= \left(\frac{\lambda(x)}{\int[\lambda_o(x)+\lambda(x)]\,dx}\right)\Big/\left(\frac{\lambda_o(x)+\lambda(x)}{\int[\lambda_o(x)+\lambda(x)]\,dx}\right) \\ &= \frac{\lambda(x)}{\lambda_o(x)+\lambda(x)} \\ &= \frac{\rho\lambda_o(x)\,f(x-x^0;\theta)}{\lambda_o(x)+\rho\lambda_o(x)\,f(x-x^0;\theta)} \\ &= \frac{\rho f(x-x^0;\theta)}{1+\rho f(x-x^0;\theta)} \end{aligned}$$

So in order to compute the likelihood function we substitute

$$p(x;\rho,\theta)=\frac{\rho f(x-x^0;\theta)}{1+\rho f(x-x^0;\theta)}$$

in (2.6), to obtain the log-likelihood function for (ρ,θ_1,θ_2) as

$$\begin{aligned} L(\rho,\theta) &= \sum_{i=1}^{n}\ln p(x_i;\rho,\theta)+\sum_{i=n+1}^{n+m}\ln[1-p(x_i;\rho,\theta)] \\ &= \sum_{i=1}^{n}\ln\left[\frac{\rho f(x_i-x^0;\theta)}{1+\rho f(x_i-x^0;\theta)}\right]+\sum_{i=n+1}^{n+m}\ln\left[1-\frac{\rho f(x_i-x^0;\theta)}{1+\rho f(x_i-x^0;\theta)}\right] \\ &= n\ln\rho+\sum_{i=1}^{n}\ln f\left(x_i-x^0;\theta\right)-\sum_{i=1}^{n+m}\ln\left[1+\rho f\left(x_i-x^0;\theta\right)\right] \end{aligned}$$

Using the formulas for f and g in (2.4) and (3.1) we get

$$\begin{aligned} L(\rho,\theta) &= n\ln\rho+\sum_{i=1}^{n}\ln f\left(x_i-x^0;\theta\right)-\sum_{i=1}^{n+m}\ln\left[1+\rho f\left(x_i-x^0;\theta\right)\right] \\ &= n\ln\rho+\sum_{i=1}^{n}\ln\left[\prod_{k=1}^{q}g\left(x_i-x_k^0;\theta\right)\right]-\sum_{i=1}^{n+m}\ln\left[1+\rho\prod_{k=1}^{q}g\left(x_i-x_k^0;\theta\right)\right] \\ &= n\ln\rho+\sum_{i=1}^{n}\sum_{k=1}^{q}\ln g\left(x_i-x_k^0;\theta\right)-\sum_{i=1}^{n+m}\ln\left[1+\rho\prod_{k=1}^{q}g\left(x_i-x_k^0;\theta\right)\right] \\ &= n\ln\rho+\sum_{i=1}^{n}\sum_{k=1}^{q}\ln\left\{1+\theta_1\exp\left[-\theta_2\left(x_i-x_k^0\right)'\left(x_i-x_k^0\right)\right]\right\} \\ &\quad -\sum_{i=1}^{n+m}\ln\left[1+\rho\prod_{k=1}^{q}\left\{1+\theta_1\exp\left[-\theta_2\left(x_i-x_k^0\right)'\left(x_i-x_k^0\right)\right]\right\}\right] \end{aligned}$$

B Sex Offender Registration

- Any felony or misdemeanor conviction or adjudication of any of the following statutes require registration:
- Indecent Solicitation of a Child; Sexual Exploitation of a Child;
- Soliciting for a Juvenile Prostitute; Keeping a place of Juvenile Prostitution; Patronizing a Juvenile Prostitute; Juvenile Pimping;
- Exploitation of a Child;
- Child Pornography;
- Criminal Sexual Assault; Aggravated Criminal Sexual Assault; Predatory Crimial Sexual Assault of a Child;
- Criminal Sexual Abuse; Aggravated Criminal Sexual Abuse;
- Ritualized Abuse of a Child;
- Forcible Detention, if the victim is under age 18;
- Indecent Solicitation of an Adult;
- Soliciting for a Prostitute, if the victim is under age 18;
- Pandering, if the victim is under age 18;
- Patronizing, if the victim is under age 18;
- Pimping, if the victim is under age 18;
- Public Indecency for a third or subsequent conviction;
- Custodial Sexual Misconduct (if convicted on or after August 22, 2002);
- Permitting Sexual Abuse of a Child;
- Kidnapping, if the victim is under age 18 and the defendant is not a parent of the victim; Aggravated Kidnapping, if the victim is under age 18 and defendant is not the parent of the victim;
- Unlawful Restraint, if the victim is under age 18 and the defendant is not the parent of the victim; Aggravated Unlawful Restraint, if the victim is under age 18 and the defendant is not the parent of the victim;
- Child Abduction by luring a child under 16 into a vehicle or building;
- First Degree Murder of a Child, victim under age 18; or
- Any attempts to commit any of the offenses listed above.

Other Qualifying Criteria for registration are:

- The offender is found not guilty by reason of insanity;
- The offender is the subject of a finding not resulting in an aquittal;
- A conviction or adjudication for a violation of federal law, the law of another state, the Uniform Code of Military Justice, or a foreign country law that is substantially equivalent to the offenses listed above;
- A juvenile is adjudicated delinquent for any of the offenses listed above; or
- A person is adjudicated as being Sexually Dangerous or Sexually Violent.

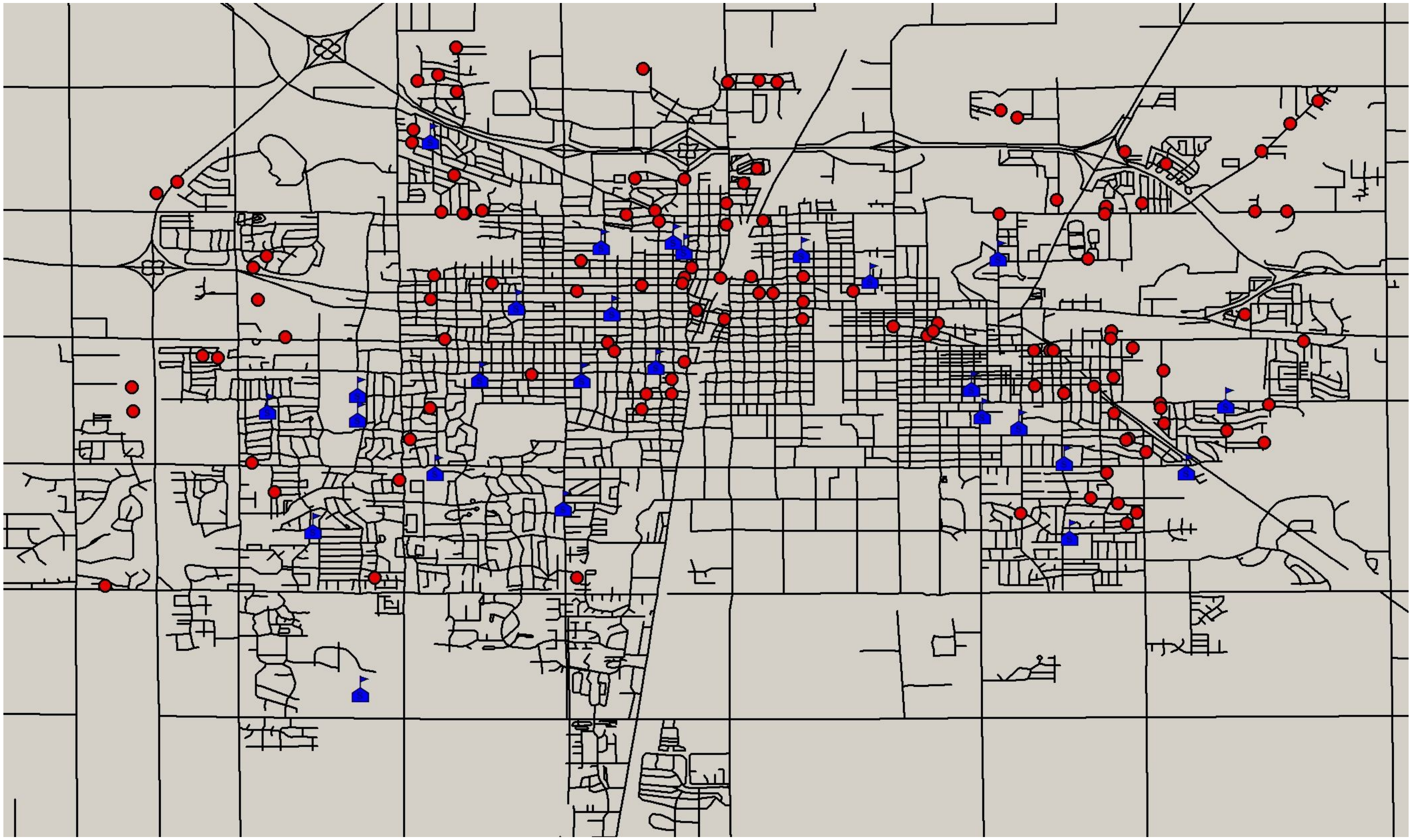


Figure 1: Residential addresses of sex offenders (red) and schools (blue) in Urbana-Champaign, IL.

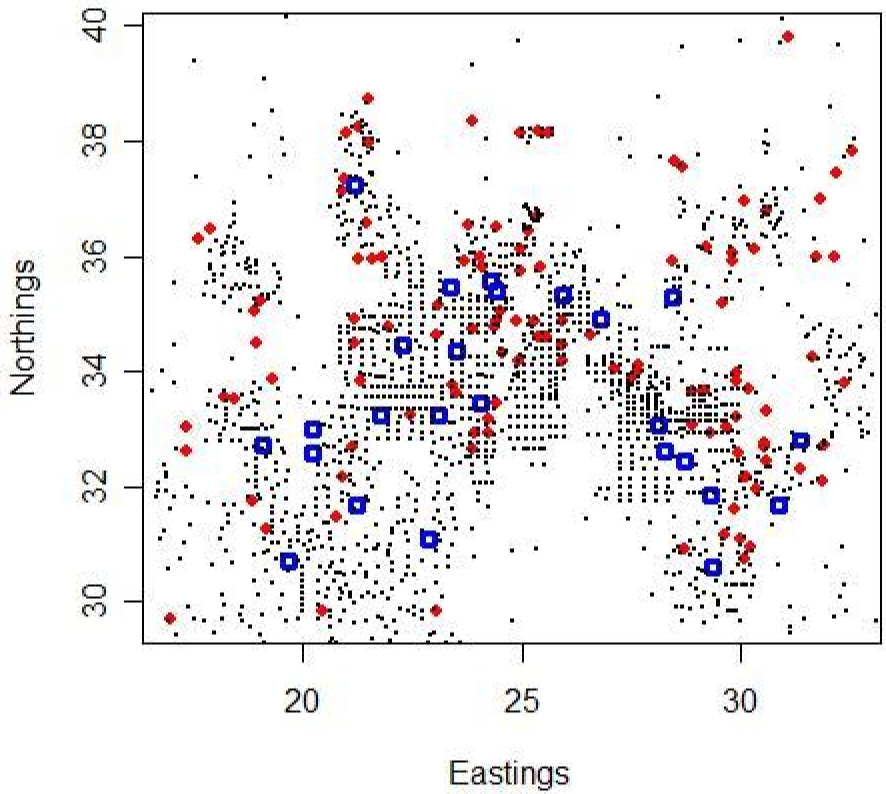


Figure 2: Projected Locations: sex offenders (red), schools (blue squares), Census block centroids (black). The distances are measured in kilometers.

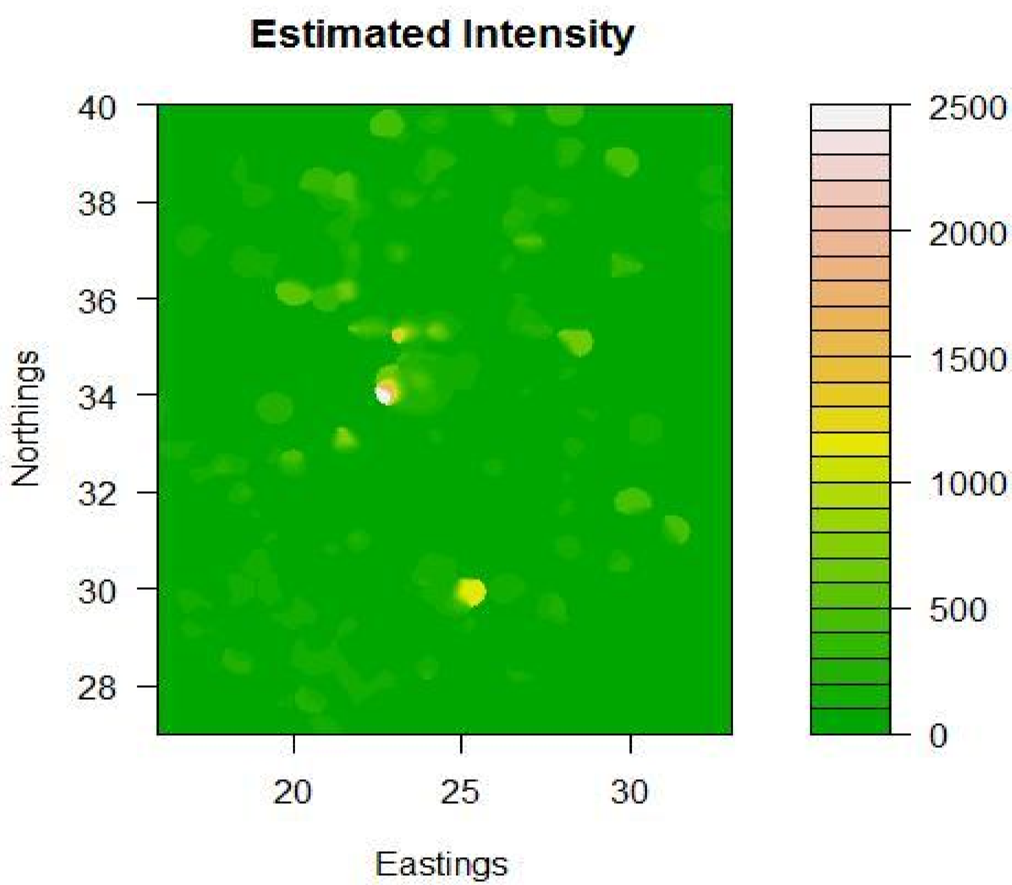


Figure 3: Estimated intensity function of the process. The distances are measured in kilometers.

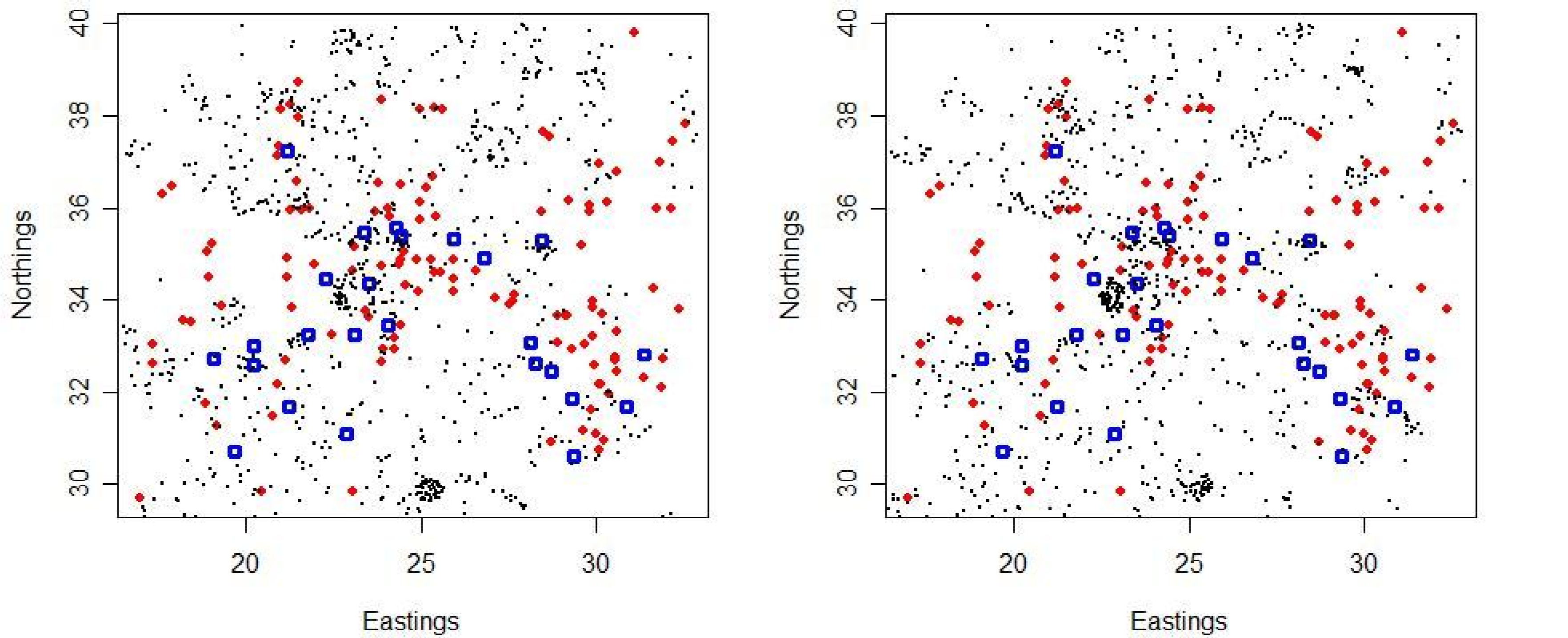
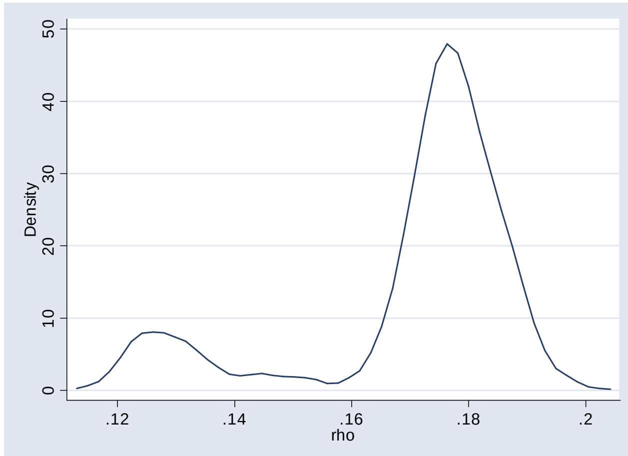
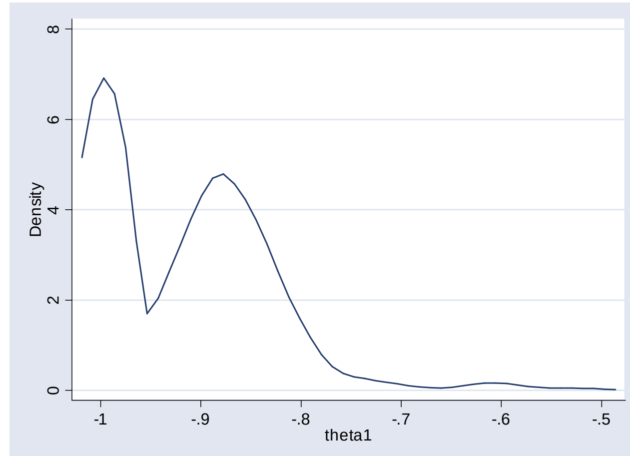


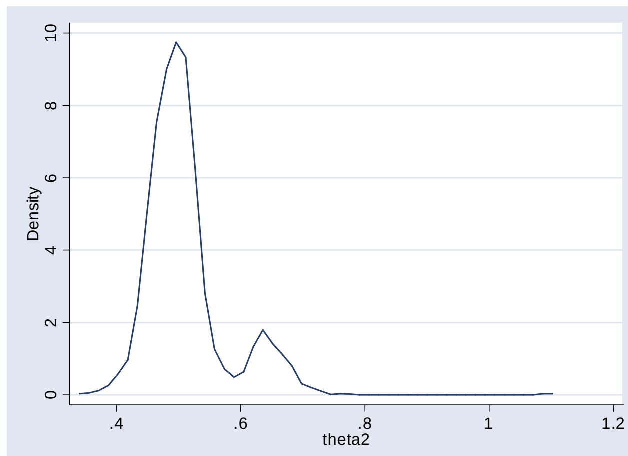
Figure 4: Two simulated samples: sex offenders (red), schools (blue squares), simulated residential locations (black). The distances are measured in kilometers.



(a)

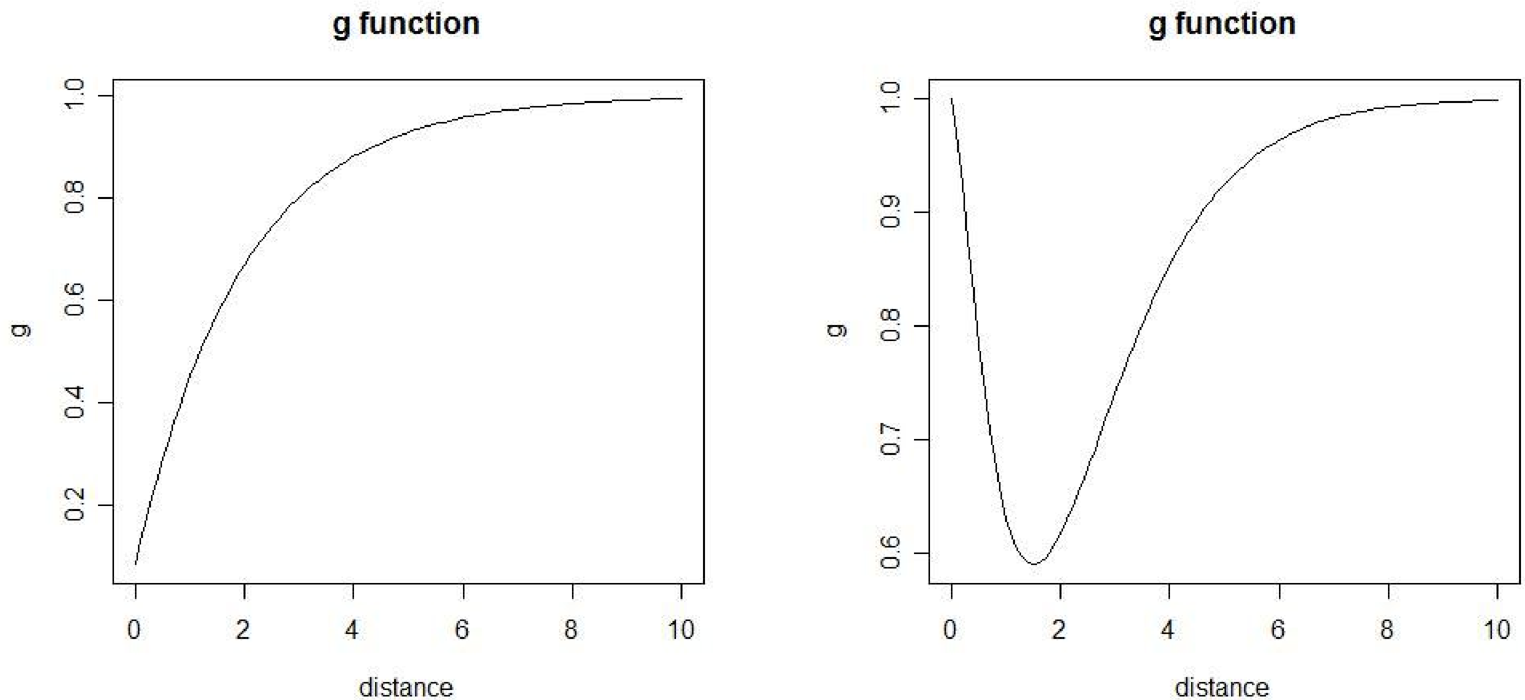


(b)



(c)

Figure 5: Empirical distribution of parameter estimates



(a) (b)

Figure 6: Estimated g functions. Distance is measured in kilometers

Table 1: Estimated Coefficients

Panel A: Baseline specification

	Coefficients				*Correlations*	
	mean	std. err.	t-stat	pvalue	θ_1	θ_2
ρ	.171	.018	9.365	.0000	.3468	−.8825
θ_1	−.916	.083	−11.050	.0000		−.6253
θ_2	.510	.065	7.869	.0000		
Proportion of Rejection $H_0 = .995$						

Panel B: Alternative Specification

	Coefficients				*Correlations*	
	mean	std. err.	t-stat	pvalue	θ_1	θ_2
ρ	.169	.004	37.597	.0000	.1163	.9692
θ_1	−.999	.002	−638.111	.0000		.1072
θ_2	1.499	.052	28.838	.0000		
Proportion of Rejection $H_0 = 1$						

Table 2: Estimated Coefficients, by school type

PANEL A: ELEMENTARY SCHOOLS

	Coefficients				*Correlations*	
	mean	std. err.	t	pval	θ_1	θ_2
ρ	.239	.107	2.234	.0329	.4797	−.2959
θ_1	−.502	.482	−1.042	.2316		−.0399
θ_2	.371	.864	.430	.3636		

Proportion of Rejections = .932

PANEL B: MIDDLE SCHOOLS

	Coefficients				*Correlations*	
	mean	std.err.	t-stat	pval	θ_1	θ_2
ρ	.184	.053	3.440	.0011	−2274	−.5678
θ_1	−.940	.276	−3.404	.0012		.1779
θ_2	2.036	3.608	.5642	.3401		

Proportion of Rejection H_0 = .780

PANEL C: HIGH SCHOOLS

	Coefficients				*Correlations*	
	mean	std. err.	t-stat	pval	θ_1	θ_2
ρ	.124	.004	34.288	.0000	−.0389	−.1353
θ_1	−.015	1.468	−.010	.3988		.2594
θ_2	8.233	3.258	2.527	.0165		

Proportion of Rejection H_0 = .070

Table 3: Estimated Coefficients, Elementary Schools

	Coefficients					
	mean	std. err.	t	pval	F	pval
ρ	.210	.070	2.987	.0047		
$\theta_{1,1}$	1.671	2.513	.665	.3197	3106.8932	.0000
$\theta_{2,1}$	13.547	5.532	2.488	.0200		
$\theta_{1,2}$	6.914	6.430	1.075	.2237	850.4606	.0000
$\theta_{2,2}$	5.903	5.716	1.033	.2339		
$\theta_{1,3}$	7.365	6.239	1.180	.1987	2257.0349	.0000
$\theta_{2,3}$	10.412	5.319	1.957	.0588		
$\theta_{1,4}$	.850	2.957	.287	.3827	3711.2671	.0000
$\theta_{2,4}$	13.151	5.020	2.619	.0130		
$\theta_{1,5}$	5.331	5.657	.942	.2557	1674.7374	.0000
$\theta_{2,5}$	10.366	5.765	1.797	.0792		
$\theta_{1,6}$	6.214	5.364	1.158	.2039	1449.0340	.0000
$\theta_{2,6}$	8.204	5.629	1.457	.1379		
$\theta_{1,7}$	7.116	5.882	1.210	.1918	5628.0316	.0000
$\theta_{2,7}$	14.945	4.502	3.319	.0016		
$\theta_{1,8}$	12.219	5.237	2.333	.0263	3741.1936	.0000
$\theta_{2,8}$	10.245	5.626	1.821	.0760		
$\theta_{1,9}$	6.220	5.873	1.059	.2276	2666.2763	.0000
$\theta_{2,9}$	11.467	5.998	1.912	.0642		
$\theta_{1,10}$	11.220	5.771	1.944	.0604	2163.7629	.0000
$\theta_{2,10}$	4.705	4.848	.970	.2490		
$\theta_{1,11}$	7.361	5.783	1.273	.1773	2017.8575	.0000
$\theta_{2,11}$	8.381	5.45	1.536	.1226		
$\theta_{1,12}$	12.179	5.418	2.248	.0320	3571.1983	.0000
$\theta_{2,12}$	9.811	5.359	1.831	.0747		
$\theta_{1,13}$	6.275	5.850	1.072	.2243	633.3144	.0000
$\theta_{2,13}$	4.013	5.062	.793	.2912		
$\theta_{1,14}$	1.841	4.137	.445	.3612	2961.5061	.0000
$\theta_{2,14}$	11.945	5.011	2.384	.0234		
$\theta_{1,15}$	4.759	5.428	.877	.2715	2447.1676	.0000
$\theta_{2,15}$	11.649	5.917	1.969	.0575		
$\theta_{1,16}$	8.667	6.291	1.378	.1544	1665.1892	.0000
$\theta_{2,16}$	7.404	5.136	1.441	.1411		
$\theta_{1,17}$	2.363	5.409	.437	.3625	5158.3229	.0000
$\theta_{2,17}$	13.959	4.954	2.818	.0076		
$\theta_{1,18}$	$-.789$	.632	-1.248	.1830	1123.9343	.0000
$\theta_{2,18}$	2.087	5.079	.411	.3665		
$\theta_{1,19}$	.993	2.380	.417	.3655	4243.2140	.0000
$\theta_{2,19}$	14.333	4.922	2.912	.0058		

Table 4: Estimated Coefficients, Middle Schools						
	Coefficients					
	mean	std. err.	t	pval	F	pval
ρ	.130	.0158	8.256	.0000		
$\theta_{1,1}$	.316	3.252	.097	.3969	703.0950	.0000
$\theta_{2,1}$	8.846	7.487	1.181	.1984		
$\theta_{1,2}$	3.220	5.832	.552	.3424	5633.5868	.0000
$\theta_{2,2}$	16.259	4.864	3.342	.0015		
$\theta_{1,3}$	3.042	5.393	.564	.3402	681.9108	.0000
$\theta_{2,3}$	8.428	7.224	1.167	.2019		
$\theta_{1,4}$	−.543	2.331	−.233	.3882	678.7872	.0000
$\theta_{2,4}$	9.357	8.714	1.074	.2240		

Table 5: Estimated Coefficients, High Schools						
	Coefficients					
	mean	std. err.	t	pval	F	pval
ρ	.1241	.002	75.954	.0000		
$\theta_{1,1}$	10.067	5.727	1.757	.0852	4585.8809	.0000
$\theta_{2,1}$	9.860	3.614	2.728	.0097		
$\theta_{1,2}$	7.408	3.516	2.107	.0434	4529.1616	.0000
$\theta_{2,2}$	9.531	3.857	2.471	.0189		
$\theta_{1,3}$	$-.505$	.644	$-.784$	.2932	690.2075	.0000
$\theta_{2,3}$	5.399	6.629	.814	.2862		