It Makes a Village:

Residential Relocation after Charter School Admission

Dr. Bartley R. Danielsen
North Carolina State University

This draft: March 28, 2010

Abstract

Although numerous studies investigate how student achievement is impacted by educational vouchers and charter schools, there appears to be no research on how school choice programs impact the surrounding environment. This study examines residential relocation of families whose children attend a charter school. I develop a theoretical model which predicts where relocating families are likely to move, given ex-ante distance and direction to the school. The model is parameterized using data from student mailing address changes. I find that families are almost twice as likely to relocate toward the school than could be expected if the school did not exert any attraction. This result has important implications for mitigating urban sprawl, fostering urban renewal, promoting sustainable real estate development, and increasing inner-city socio-economic and racial diversity.

Dr. Bartley R. Danielsen
North Carolina State University
Raleigh, NC 27695-7229

919.513.0003 (v)
brdaniel@ncsu.edu
I. Introduction

Baum (2004) argues that in any effort to manage urban sprawl, a focus on education is a quintessential issue. Inner-city schools have become increasingly racially segregated as their performance has lagged behind suburban counterparts. As inner-city schools deteriorated, middle class families fled to suburban locations. The fallout from this middle-class migration includes increased urban sprawl, concentrated inner-city poverty, and reduced inner-city racial diversity.

Baum proposes as one solution to inner-city school dysfunction that students be removed from poor-performing city schools and sent to good schools in the suburbs. In fact, this approach has been pursued with some success in Boston and St. Louis. Both cities previously bused inner-city students to suburban schools in order to comply with judicial integration mandates. A third program, also generally viewed as an educational success, is the Gautreaux program in Chicago. The Gautreaux program provided housing vouchers to the urban poor that could be used in the suburbs.

The Gautreaux program resulted from litigation against the Chicago Housing Authority (CHA) and the Department of Housing and Urban Development (HUD). The CHA and HUD previously racially segregated black families in Chicago in violation of civil rights laws. The Gautreaux litigation settlement included relocating black families to majority-white suburbs, and participants generally reported a preference for the safer neighborhoods and higher quality schools in the suburbs. In 1998, when Gautreaux met its’ target of relocating 7100 families, the program terminated, but the success of the program led to similar HUD programs that remained an integral part of federal housing policy.

An ironic element of Baum’s proposal to educate inner-city children in suburban schools is that bussing and relocating inner-city students to the suburbs only exacerbates urban sprawl, even as it attempts to mitigate prior sprawl’s unpleasant outcomes. At its core, the policy proposes to send more
students and families from the urban core to the suburbs, although Baum acknowledges that “in the long-run there is no alternative to making city schools work.” Among the reform efforts considered to make city schools work are charter schools and other school choice schemes.

Where schools of choice are concerned, Baum appears to be of two minds. He asserts that “establishment of reliable schools in working-class and middle class city neighborhoods would hold remaining families with resources and draw back” suburban families. While this assertion sounds plausible (and hopeful), there appears to be no empirical support for this assertion. In fact, Baum acknowledges that “private schools to which middle-class and elite families send their children are rarely neighborhood schools,” and “low-income voucher programs give up neighborhood schools for better schools.” If this is true, it remains an open question whether charter schools or voucher programs can be effective at reducing urban sprawl by evolving into neighborhood schools. Although it seems plausible that middle-class families will chose to live near such schools, empirical evidence is lacking.

This study is an effort to initiate inquiry regarding the movement patterns of families who attend charter schools, magnet schools, and private schools. I believe that this is the first study investigating migration patterns of families who are able to exercise educational school choice. Families who have a choice of which school their child attends, but who are freed of the need to relocate into a specific school district, may choose to relocate in two distinct patterns that can have environmental, urban planning, real estate development, transportation, and urban sprawl ramifications. On one hand, families may choose to live closer to the schools that their children attend. On the other, being no longer geographically tied to the school by assignment policy, families may choose to move farther away from the school.

Previous research on housing location choices, as they relate to adult employment location, lends support to the hypothesis that, families may choose to live closer to schools that their children attend. Clark, Huang and Withers (2003) observe that people tend to relocate closer to their work locations when they move. Two-worker families consider the commutes of both parties when choosing to relocate. Interestingly, in many instances, two-earner households are more likely to move closer to the wife’s workplace than the husband’s. This pattern may be attributable to females’ greater need to balance the
dual role of mother and worker. A similar logic would suggest that home-to-school commutes are likely to be an important relocation driver.

Empirical evidence that families move toward their children’s schools when relocation is not compelled by a “neighborhood schools” policy, would suggest that urban school choice programs may produce subsequent relocation behaviors that decrease urban sprawl, reduce environmental impacts (carbon footprints), and increase inner-city diversity.

However, school choice may have an alternative impact on family housing choices as it relates to urban sprawl. Geographically based school assignments usually require families to live in close proximity to the schools that they attend. Freed from an explicit geographic tether, families with students attending schools of their own choice might decide to move farther from the school and increase the school commute. Consider, for example, the work-commuting choices that people make. We know that households work-commuting distance is not the only reason for residential relocation. Accessibility to work is only one variable. In fact, it appears that many moves within the city, in effect, hold the distance to work as a constant. Previous work describes an ‘indifference zone’ within which commuters are relatively indifferent to access to work (Getis, 1969). Brown (1975) found that households with employment changes outside their original work zone were much more likely to move than were households within the original work zone. In fact, there appears to be a marked tendency for households to move closer to their workplace as the ex-ante separation increases. Simply put, if a household is a long distance from the workplace, when the household moves it is likely to move nearer the workplace. Thus, we might expect that home-to-school commutes will follow a similar dynamic. Even if families who live far from a school tend to move closer, those who live near the school may choose to live further away, simply because their housing choice no longer predetermines school enrollment.

Taken together, we can expect that families who have freedom of choice that is disentangled from geographic eligibility will occasionally move closer to the school, and occasionally move away. It is an open question as to which choice will dominate. The resolution of this issue is critical to assessing
whether school choice will increase urban sprawl and accelerate inner-city decline, or decrease sprawl and improve inner city schools.

A literature review reveals no prior research concerning how school choice programs impact family relocation or the environment. However, there is a literature that explores how commuters relocate relative to the workplace. This study follows the framework developed in that literature.

II. Data, Hypothesis and Descriptive Interpretations

The data used to conduct this analysis is provided by a charter school in the Raleigh-Durham, North Carolina area. By state law, admission to charter schools is conducted by lottery. Because North Carolina caps the number of state charter schools at 100, it is not uncommon for the demand for charter schools to exceed the available seats. This is the case for the school in question. Application to the school entitles the applicant to participate in the lottery, but there is no guarantee that the student will be admitted.

Preference is granted to applicants who have a sibling already enrolled at the school. If there are more seats available in the class than the number of sibling students who apply, all of the sibling students are admitted, and a lottery is held for the remaining seats. If there are fewer seats than sibling students, the lottery is held for the sibling students only, and no outside applicants are admitted.

Each applicant must complete an application containing, among other data, the mailing address of the applicant’s family. Once the student is admitted, this application is retained in the student’s permanent record. Using the permanent record files, we have assembled the initial mailing addresses for all students attending the school. The school continuously updates student mailing addresses for general purposes, and by comparing the address on each student’s application to his/her subsequent mailing list address, we are able to determine which students have moved since being admitted to the school. In addition, by matching the last names and mailing addresses for students, we are able to determine which students are members of single-family. Moreover, we can identify which student in any family as the first admitted sibling. The data on these students is of interest in this research.
Because siblings are granted priority admission, a family who has one student admitted to the school can expect that siblings will gain admission in a later year. This may be important for families with multiple children, even if younger children are not yet of school age. Enrolling a child in the school creates a pathway for enrolling all other school-aged siblings once they are ready to attend the school. In other words, the family secures the right for each child to attend the school once the first child is admitted. Thus, admission of the first child to the school confers a valuable right which may impact the family’s residential location choice.

With this in mind, admission of the family's first child would appear to be a triggering event most likely to alter a family’s optimal residential location choice. Therefore, for each family, we identify both the family residence location prior to the first admission to the school, and the subsequent mailing address as of January 2009.

The data that we have collected reveals 662 families had at least one student attending the school. The application mailing addresses for the first-admitted child in four instances cannot be ascribed to a true place of residence because a Post Office Box is given. The other addresses described are presumed to be true residential addresses.

We geocoded each address using the ArcGIS 9.2 "Geocode Addresses Tool," which utilizes street centerline data for address ranges. Specifically, we used the North Carolina Department of Transportation's 2007 Integrated Statewide Road Network database. 1 We also geocode the school’s location. The result of the geocoding is a shapefile of points, with each point representing the address location (longitude and latitude) for a single record in the data table. Any addresses that did not properly geocode had points created based upon manual searches using Mapquest and Google Earth. The attribute data table of each point contained the record ID, student address, and geographic latitude/longitude coordinates.

Using the January 2009 mailing addresses of all students, we identified 178 families that had moved since the first child applied to the school. For these new addresses, we repeated the process and geocoded the new addresses.

Finally, we use Hawth’s Tools, an ArcGIS 9.2 extension, to calculate the linear distance from each address to the school. We also calculate bearing and turn angle metrics which are discussed later in the paper. Hawth’s Tools are designed specifically for ecology related analyses such as this. We also access Google maps to calculate the nonlinear road-commuting distance and the estimated commuting time from each address to the school.

We expect that family relocation decisions are likely to be determined by commuting time and distance rather than linear distance. However, the model which we construct later in the paper uses trigonometric functions that presume linear movements. In order to obtain some comfort that linear distance provides a reasonable proxy for families’ more likely decision variables of nonlinear road commute and commuting time, we have calculated the correlation between each of these three measures. These correlations are presented in Table 1.

### Table 1:
**Correlation of linear distance, drive distance, and drive time for accepted applicants (first in family)**

<table>
<thead>
<tr>
<th></th>
<th>Linear distance</th>
<th>Drive distance</th>
<th>Drive Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear distance</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drive distance</td>
<td>0.9901</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Drive time</td>
<td>0.9554</td>
<td>0.9638</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Notice that all of these variables are very highly correlated. In particular, the drive distance is highly correlated with the linear distance. The very-high level of correlation appears to be related to the fact that the school location is common to each commute. Given that the last leg of the commute follows the same few paths for all commuters, the linear distance maps very closely to the drive distance.

---

2 The Hawth’s Tool module used is “Calculate Movement Parameters”. Documentation can be found at the following: [http://www.spatailecology.com/htools/moveparamssimple.php](http://www.spatailecology.com/htools/moveparamssimple.php)
We are able to identify a residential address at the time of application for 658 of the 662 families admitted to the school. Table 2 provides descriptive statistics concerning linear distance, in miles, from each family’s original address to the school’s location. Admitted applicants, on average, lived 5.77 miles from the school, and the median distance from the school was 4.59 miles. Less than one percent of the admitted students lived within a quarter of a mile from the school. Approximately 95% live within 15 miles.

Table 2:  
Original linear distance in miles from home to school for accepted applicants (first in family)

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Smallest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>.2611511</td>
</tr>
<tr>
<td>5%</td>
<td>.920028</td>
</tr>
<tr>
<td>10%</td>
<td>1.510568</td>
</tr>
<tr>
<td>25%</td>
<td>2.572471</td>
</tr>
<tr>
<td>50%</td>
<td>4.599287</td>
</tr>
<tr>
<td>75%</td>
<td>7.245863</td>
</tr>
<tr>
<td>90%</td>
<td>10.39763</td>
</tr>
<tr>
<td>95%</td>
<td>14.34032</td>
</tr>
<tr>
<td>99%</td>
<td>24.82562</td>
</tr>
</tbody>
</table>

Comparing the application addresses to the subsequent mailing addresses, we find that 188 of the families (26.57%) changed addresses after they were admitted to the school. 470 of the families did not change address. We assume that a change of mailing address constitutes a change of residence, but this need not be the case. For instance, a family might use a business address or a post office address for receiving personal correspondence. In that case, the change will be misinterpreted as a change of residence. School administrators also point out that a small number of students have divorced parents with joint custody. We cannot systematically identify these students, and we have no means of determining what impact these family arrangements might have on the data. In any event, noise that is introduced by these factors should bias against finding school commute to be an important factor in relocation decisions.
Assuming that families make relocation decisions on the basis of commute time, we might expect that families who live a long distance from the school would be more likely to relocate. To test this hypothesis, we specify the following probit model:

\[ P(\text{Moved}_i = 1) = \alpha_0 + \alpha_1 \text{Distance}_i + \alpha_2 \text{Years}_i + \varepsilon_i \]

Where \((\text{Moved}_i = 1)\) indicated that a family \(i\) moved after admission, and \((\text{Moved}_i = 0)\) indicates that the family did not move. Distance\(_i\) is the pre-move linear distance from the school, and Years\(_i\) is the number of years that the student has been enrolled at the school. Students who have attended the school for a longer period of time, are more likely to have moved, without regard to motivation. Thus, the coefficient \(\alpha_2\) should be positive. However, the coefficient of interest is \(\alpha_1\). If families with longer home-to-school commutes move to reduce the commute, we expect those families with longer commutes to be more likely to move. If this is true, \(\alpha_1\) will be positive.

The first regression specification in Table 3 presents the results of the hypothesized model. T-statistics for each coefficient estimate are given in parentheses.

### Table 3:
**Probit estimation of relocation probability given distance to school**

<table>
<thead>
<tr>
<th></th>
<th>Moved (one if family moved, zero otherwise)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.01833</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
</tr>
<tr>
<td>Distance</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>(1.96)</td>
</tr>
<tr>
<td>Years</td>
<td>0.05158</td>
</tr>
<tr>
<td></td>
<td>(8.75)</td>
</tr>
<tr>
<td>Admitted Grade</td>
<td>-0.0516</td>
</tr>
<tr>
<td></td>
<td>(-7.68)</td>
</tr>
<tr>
<td>Current Grade</td>
<td>0.0515</td>
</tr>
<tr>
<td></td>
<td>(8.56)</td>
</tr>
<tr>
<td>(N)</td>
<td>650</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.1096</td>
</tr>
<tr>
<td>(F(\text{obs}, \text{N-obs}))</td>
<td>41.35</td>
</tr>
<tr>
<td>Prob. &gt; (F)</td>
<td>0</td>
</tr>
</tbody>
</table>
As expected, the number of years that the student has attended the charter school is highly correlated with the probability of a move. Obviously, the more time that has elapsed between the two observation points, the more likely it is that a move will have occurred. More importantly, the distance that the family originally commuted to school is also positively correlated with the move probability. This is consistent with the hypothesis that families will move closer to the school once they have the right to attend.

The second specification in Table 3 details decomposes the time that the student has been enrolled at the school into the student’s Current Grade and the student’s Admitted Grade. The difference between Current Grade and Admitted Grade is the value of Years in the original specification. The negative coefficient on Admitted Grade indicates that the younger the student was when he was admitted, the more likely the family was to relocate. This is consistent with families choosing to relocate when they expect that their children will be enrolled at the school for a long period of time. For families that expect to be affiliated with the school for many years, the relative benefits of moving increase. The positive coefficient on Current Grade indicates that older students are more likely to have moved since enrolling.

III. A Model of School Attraction

The foregoing frequency distributions and probit analyses are helpful in describing the relationship between school location and relocation choice. However, if we wish to fully understand the magnitude of the school’s attraction in residential relocation decisions, a theoretical model of the relocation decision is useful. Ideally, a model of school site attraction will (1) provide testable hypotheses concerning the probability of moving closer to or further from the school, and (2) provide testable hypotheses concerning the effect of distance on school site attraction.
In order to simplify exposition of the model that will follow, let us first consider a simple conceptualization of one family’s residential relocation. Figure 1 presents a vector structure of the school-residence relationships. In the Figure 1 diagram, the student lives at the residence $R_{\text{Old}}$ prior to enrolling in the school. The distance that the student lives from the school is identified as $d_O$. After being admitted to the school, the student moves to a new residence, designated as $R_{\text{New}}$. The distance moved from $R_{\text{Old}}$ to $R_{\text{New}}$ is designated as vector $X$. After moving to $R_{\text{New}}$, the new commuting distance to the school is designated by the vector $d_N$. Summarizing the distances involved in this move, the student moved $X$ miles from $R_{\text{Old}}$ to $R_{\text{New}}$, and the commute distance to the school changed from the $d_O$ to $d_N$.

**Figure 1**

![Vector Diagram](image)

In addition to the distances that have been identified, another important aspect of this conceptualization concerns the angle theta. Theta is the angle formed by moving from vector $d_O$ to vector $X$. If a student moved directly toward the school, the value of theta would be zero. For movements in a counter-clockwise direction from the original school bearing, the value of theta is between $-\pi$ and zero ($-\pi < \theta < 0$). In the Figure 1 example, the value of $\theta$ would be approximately $-\pi/4$, corresponding to a 45 degree angle moving counter-clockwise. Similarly,
for movements in a clockwise direction from the original school bearing, the value of theta is between zero and π (0 < θ < π). The importance of theta will be seen in the further development of the model.

We are interested in the relationship between distances from the student’s residence before and after the move. The conceptualization of this relationship can now be structured as a model with two parameters in which each student’s move is described by the vector X, which has both a length and a direction. Thus, the distribution of these moves across the full sample is a joint distribution of directions and lengths for all X’s.

This brings us to a formal model of the relationships conceptualized in Figure 1. Quigley and Weinberg (1977), Clark and Burt (1980), and Clark, Huang and Withers (2003) consider relocations as a function of move distances from workplaces (analogous to this study of moves related to school location). These papers make the empirical observation that move distances are distributed exponentially:

\[ f(X; \alpha) = \alpha e^{-\alpha X}, \quad X > 0 \]  
(1)

where X is the distance in miles. Here \( \alpha \) is the rate parameter of the distribution, and the distribution is supported on the interval \([0, \infty)\).

Foreshadowing later results, we will find it useful to also provide a more general distributional assumption than the exponential distribution. The exponential distribution is a special case of the more general gamma distribution,

\[ g(X; \varphi, \alpha) = \frac{\alpha^\varphi}{\Gamma(\varphi)} X^{\varphi-1} e^{-\alpha X}, \quad X > 0 \text{ and } \varphi, \alpha > 0. \]  
(2)

This gamma distribution is parameterized in terms of a shape parameter \( \varphi \), as well as the rate parameter \( \alpha \). The function \( \Gamma(\varphi) \) is defined to satisfy \( \Gamma(\varphi) = (\varphi - 1)! \) for all positive integers \( \varphi \), and
to smoothly interpolate the factorial between integers. For the special case of the shape parameter $\varphi=1$, the gamma distribution in equation 2 becomes the exponential distribution of equation 1.

A second assumption of our model is that the move directions for students follow a von Mises distribution (Gaile and Burt, 1976). The von Mises distribution is also known as the circular normal distribution. Accordingly, it can be viewed as an analogue to the normal distribution that is useful for analyzing two-dimensional data. The parameters of the von Mises distribution are $\mu$ and $\kappa$, which are analogous to the normal distribution’s parameters $\mu$ and $\sigma^2$. Actually, $\kappa$ is analogous to the inverse of $\sigma^2$, $(1/\sigma^2)$.

The assumption that student movements are, on average, in the direction of the school is captured as $\mu=0$. (This assumption is subject to subsequent testing.) For a mean direction of zero, the von Mises density function is defined as

$$
\nu(\theta) = \frac{1}{2\pi I_0(k)} e^{k \cos(\theta)}, \quad -\pi < \theta < \pi, k \geq 0
$$

(3)

where $\theta$ is the move direction described in Figure 1, measured in radians. $I_0$ is a modified Bessel function of the first kind and order zero.

As noted above, the dispersion of the von Mises is a function of $k$. Figure 2 depicts several potential values of $k$ for a distribution with mean direction of movement $\mu$.

**Figure 2**
When $k = 0$, movements are in no preferred direction, without regard to the value of $\mu$. However, as $k$ increases, the magnitude of movements in the $\mu$ direction increases. One of the tests which follows will estimate the magnitude of $k$, with $k$ serving as a measure of the attraction on the student that is exerted by the school location. The larger $k$ is, the stronger the relocation attraction of the school.

We also assume that move directions and distances are independent of one another. This assumption aids tractability but biases against finding confirming empirical support if the assumption is invalid. Thus, as noted by Clark, Huang and Withers (2003) “if the fit between observed and expected is good, we are confident of the results of the model.” Accordingly, the joint probability distribution of movement distance and direction is described by

$$c(X, \theta) = g(X) \nu(\theta)$$

(3)

Given these assumptions we develop a model of the likelihood that a student will move into a particular area defined by two distances ($X_1$ and $X_2$) and two angles ($\theta_1$ and $\theta_2$),

$$P(X_1 < X < X_2, \theta_1 < \theta < \theta_2) = \int_{X_1}^{X_2} \int_{\theta_1}^{\theta_2} c(X, \theta) d\theta dX$$

(4)

where

$$c(X, \theta) = g(X) \nu(\theta) = \left(\frac{\alpha^\nu}{\Gamma(\nu)}X^{\nu-1}e^{-\alpha X}\right)\left(\frac{1}{2\pi\sigma^2}\right) e^{k \cos(\theta)}$$

Recall from Figure 1 that students move closer to the school for $d_N < d_O$. Thus, we are specifically interested in the region where $d_N < d_O$. Specifically, we wish to solve for $P(d_N < d_O)$. From the law of cosines

$$(d_N)^2 = (d_O)^2 + (X)^2 - 2(d_O X) \cos \theta$$

(5)

Thus,

$$P(d_N < d_O) = P((d_N)^2 < (d_O)^2)$$
\[
= P \left( (d_0)^2 + (X)^2 - 2(d_0 X) \cos \theta < (d_0)^2 \right) \\
= P \left( X < 2(d_0) \cos \theta \right) \\
= \int_{\pi/2}^{\pi/2} \int_{0}^{2(d_0) \cos \theta} c(X, \theta) dX \, d\theta 
\]

\[
P(d_N < d_0) = 2 \int_{0}^{\pi/2} \int_{0}^{2d_0 \cos \theta} c(x, \theta) x dx \, d\theta 
\]

\[
= 2 \int_{0}^{\pi/2} \int_{0}^{2d_0 \cos \theta} \left( \frac{\alpha^\phi}{\Gamma(\varphi)} x^{\varphi-1} e^{-\alpha x} \right) \left( \frac{1}{2\pi I_0(k)} e^{k \cos \theta} \right) x dx \, d\theta 
\]

\[
= \frac{\alpha^\phi}{\pi I_0(k) \Gamma(\varphi)} \int_{0}^{\pi/2} e^{k \cos \theta} \int_{0}^{2d_0 \cos \theta} x^{\varphi-1} e^{-\alpha x} x dx \, d\theta 
\]

Let \( t = \cos \theta, \, dt = d\cos \theta = -\sin \theta \, d\theta. \)

Because \( \cos^2 \theta + \sin^2 \theta = 1, \) \( d\theta = \frac{1}{-\sin \theta} dt = \frac{1}{\sqrt{1-t^2}} \, dt. \)

\[
P(d_N < d_0) = \frac{\alpha^\phi}{\pi I_0(k) \Gamma(\varphi)} \int_{0}^{1} \frac{1}{\sqrt{1-t^2}} e^{kt} \int_{0}^{2d_0 t} x^{\varphi-1} e^{-ax} x dx \, dt 
\]

Equation 7 can be evaluated for various values of \( k \) and \( d_0 \) using numerical integration.

This allows us to establish the relationship between \( P(d_N < d_0) \) and \( d_0. \)

**IV. Tests of School Attraction**

The tests conducted in this subsection take place in two parts. The first test assesses the distributions of observed and expected move distance. The mean move distance is 7.585 miles, and the estimate of exponential rate parameter \( (\alpha = 1/\text{mean distance moved}) \) is 0.132.
Although previous research has concluded that residential relocation follows an exponential distribution, we should consider whether this distribution fits our particular sample. Much prior research examines movements between zip codes of census groups or blocks. Thus, no moves of very short distances are resident in the data. By necessity, under those circumstances, fitting to the exponential distribution (or any distribution) is conducted without the benefit of very-short-distance moves. We have finer data in that we use actual street address coordinates. Although the mean move distance is 7.585 miles, a quarter the moves observed are less than 2.66 miles. Thus, we are in a position to fit the data more precisely prior researchers have been able to do.

The Kolmogorov–Smirnov statistic quantifies the distance between the empirical data and the hypothesized cumulative distribution function. Testing the null hypothesis that the sample is drawn from an exponential distribution with rate parameter $\alpha$ equal to 0.132, The Kolmogorov-Smirnov Goodness-of-Fit test yields a p-value = 0.009. Therefore we must reject that the exponential is the appropriate distribution to use for mean distance moved. Thus, we also consider the more flexible gamma distribution. The exponential distribution is a constrained form of the gamma distribution; the shape parameter is equal to 1.

Assuming that the observed moves are drawn from the gamma distribution, we find parameter estimates of $\alpha = 0.1688766$, and shape parameter $\phi = 1.280951$. Figure 3 plots the fitted exponential density function and the fitted gamma density function against the move distance. The mean of the gamma distribution, $(\alpha^{-1})(\phi)$, is also 7.585 miles.
The move distance conforms to the X vector in Figure 1, and the X value in the theoretical distributions from Equation 1 and Equation 2. Notice that the exponential function (dashed curve) has a modal value of zero. This distribution would suggest that the most likely move distance for any residential relocation is to the nearest alternative residence – a move to the house next door. The fitted gamma distribution (the solid curve) produces a modal move 1.7 miles from the original location. This seems to be a reasonable finding. Rather than changing homes within the same neighborhood, the gamma function suggests that relocaters are more likely to move to a nearby neighborhood than immediately next door.

The Kolmogorov-Smirnov Goodness-of-Fit test for the gamma function yields a p-value = 0.356, and we fail to reject the hypothesis that the move distances are drawn from this maximum-likelihood-estimated fitted distribution. Figure 4 graphically compares the goodness of fit for the exponential and
gamma distribution. Notice that the exponential function overestimates the number of short-distance moves, relative to the fitted gamma distribution.

**Figure 4**

![Exponential distribution goodness of fit](image)

![Gamma distribution goodness of fit](image)

Turning to our tests of move direction, the direction of each move in the sample can be represented by a vector with direction $\theta$ whose length is one (unit vector). The use of unit vectors conforms to the theoretical assumption that move direction and move length are independent. Summing all the sample vectors results in a vector $\mathbf{R}$ where

$$\theta_R = \tan^{-1} \frac{\frac{1}{n} \sum \sin \theta_i}{\frac{1}{n} \sum \cos \theta_i}$$
is a measure of mean move direction. The length of vector $\mathbf{R}$ also reflects the extent of clustering in the sample’s mean direction. This clustering is analogous to the variance in non-directional data. Standardizing by the number of observations in the sample yields an index $\bar{R}$ with a value between zero and one.

$$\bar{R} = \frac{\mathbf{R}}{n} = \frac{\sqrt{\left(\sum \sin \theta_i\right)^2 + \left(\sum \cos \theta_i\right)^2}}{n}$$

$\bar{R}$ is a function of the concentration parameter $k$ by virtue of

$$\bar{R} = \frac{I_1(k)}{I_0(k)}$$

where $I_0(k)$ is a modified Basel function of the first kind and zero order.

For the sample of relocating families in the current study, $\theta_R$ equals 0.147 radians, or 8.421 degrees. The clustering index $\bar{R}$ equals 0.5225, yielding concentration parameter $k = 1.120$.

For the von Mises distribution parent population when $n$ is large and $k = 0$ the statistic $2n\bar{R}^2$ is approximately $\chi^2$ distributed with two degrees of freedom.\(^3\) For the current test, the value is 102.667 which is far above any reasonable cutoff value ($p=0.05$, cutoff value=5.99). Thus, we reject the null hypothesis of $k=0$ (no bias).

Given a move direction bias, we test the assumption that the move directions are biased toward the school. This test assumes the school is the attractor and tests whether or not we can reject that assumption. The 95% confidence interval around the school direction can be written as $0 \pm 1.96/\sqrt{nk\bar{R}}$

$$= 0 \pm 1.96/\sqrt{(188)(1.120)(0.5225)} = 0 \pm 0.1790 \text{ radians.}$$

Because $-0.1790 < \theta_R < 0.1790$, we accept the hypothesis that the move directions are concentrated toward the school.

As a point of reference, previous studies by Clark and Burt (1980) and Clark, Huang and Withers (2003) consider workplace attraction. The first paper studied workplace attraction in the Milwaukee

\(^3\) See Mardia (1972).
metropolitan area. This study found a concentration parameter $k=0.638$. The second study conducted similar tests to gauge Seattle area work-place attraction and yielded a parameter $k=0.668$. Tests on a subset of female-worker relocations yielded $k=0.831$. Notice that the value associated here with school attraction ($k=1.120$) is significantly larger than previously reported work-place attraction measures.

A. Imputed Probabilities of Toward-School Migration

Conditional upon a family moving, we are interested on assessing the probability that it will move toward the school. Figure 5 provides a simple graphic representation of the question. Given that the family’s original home, $R_{\text{Old}}$ is a distance $d_{\text{o}}$ from the school, we are interested in the probabilities that the family will move to a location that is closer to school - the shaded area in the figure.

Figure 5

1. Base case probabilities ($k=0$)

There is some probability that the family would move closer to the school even if the school were not a relocation attractor. This is the probability when $k=0$. To obtain this baseline probability, we numerically solve equation 7 for various values of $d_{\text{o}}$, given $k = 0$, $\alpha = 0.1688766$ and $\phi = 1.280951$. 
For families already living near the school, the probability that they will move close is small. A family living a mile from the school only has a 0.055 probability of moving closer. However, for $d_o=10$, the probability of moving closer rises to 0.367. In the limit, the probability rises to 50%.

2. **Imputed move probabilities (k=1.120)**

Reassessing the probability that a family will move closer, given the observed attraction exerted by the school, we reevaluate equation 7 for values of $d_o$, given $k = 1.120$. The parameters $\alpha$ and $\phi$ are unchanged. For families already living a mile from the school, the probability of moving closer nearly doubles, rising from 0.055 to 0.106. For $d_o=10$, the probability rises from 0.367 to 0.669.

Although the increase in probability can be estimated for longer initial commutes, only 8% of the movers had initial commutes of over 15 miles. With relatively few actual observations, we are not confident that the imputed probabilities would be meaningful for extreme values of $d_o$. For example, if we fit the model for $d_o = 100$, $P(d_N < d_o) = 0.815$, but no initial commutes were this long, and it seems likely that the parameters of the fitted gamma distribution would be altered if an observed $d_o$ had existed.

Figure 6 provides a visual depiction of the increase in $P(d_N < d_o)$ for $1 \leq d_o \leq 15$. The increase in Figure 6 is defined as $P(d_N < d_o)|_{k=1.120} - P(d_N < d_o)|_{k=0}$.
Figure 7 depicts the ratio of \( P(d_N < d_O)\rvert_{k=1.120} \) to \( P(d_N < d_O)\rvert_{k=0} \) for \( 1 \leq d_O \leq 15 \). As noted above, for families already living a mile from the school, the probability of moving closer nearly doubles. Even for families living 15 miles away from the school, the probability of moving closer is almost 1.8 times greater.

![Figure 7](image)

**Figure 7**

*Increase in the probability of moving closer to the school*

\[
\frac{P(d_N < d_O)\rvert_{k=1.120}}{P(d_N < d_O)\rvert_{k=0}}
\]

---

**B. Further Analysis**

To help the reader more clearly visualize the move pattern of relocating families, we present a rose diagram in Figure 8. The rose diagram aggregates moves that occur in common directions into several bins. The rose diagram resembles a pie chart, except that each bin (sector) has an equal angle. Rather than alter the central angles to account for different numbers of observations in each sector, we extend each sector from the center of the circle by varying distances to illustrate the number of moves that occur in a particular direction.

Families can move in any direction, and we have segmented the circle into twelve 30-degree bins. The right-most segment is centered on the school so that this bin contains all observations for families moving in a direction within 15 degrees of \( \theta = 0 \). In order to make the constructed areas proportional to
the frequencies, the length of each wedge is proportional to the square root of the number of observations. In this graph, the fraction of the observations represented by the largest wedge is 30.3%, and the fraction represented by the smallest wedge shown is 1.60%. In this framework, the magnitude of the family relocation bias is obvious.

**Figure 8:**
Move Directions with 12 Bins

We have also calculated the mean distance moved by the families in each bin shown in Figure 8. The mean move distances are graphically depicted in Figure 9, with the values for the mean and standard deviations shown below the figure. The group names in the legend reflect the geographic bounds on each bin. The bounds are identical to those used to construct Figure 8. The first group is for movers in the direction of the school which includes moves within 0+15 and move at an angle greater than 345 degrees. The bins in the table are listed in a counter-clockwise direction from the school.
Causal observation suggests that families moving toward the school move much farther, on average, than those moving away. The mean distance moved toward the school is 10.9 miles, and the mean distance moved directly away from the school is only 1.7 miles. (Group 165-195 is centered on 180 degrees from the school.) More rigorously, based on a small ANOVA test p-value (p = .0036), we
conclude that the distance moved is affected by the direction, and the assumption that distance and direction are independent does not hold.

Because the independence assumption is violated, the various probabilities calculated for \( k=1.12 \) are lower-bound values. The actual increase in probabilities that families relocate closer to the school must be even higher than that depicted in Figures 6 and 7.

V. Conclusion:

This study provides a theoretical foundation for considering environmental implications of school choice plans. More narrowly, I develop a model of move distance (distributed gamma) and direction (distributed von Mises) to predict family relocation choice, relative to school location. The model is parameterized using data from student mailing-address changes. The model predict that families are almost twice as likely to relocate toward the school than could be expected if the school did not exert any attraction. Because move distance and direction in the sample are not independent, the theoretical model actually underestimates the magnitude of the school’s attraction. This result has important implications for the potential role of charter schools and vouchers in mitigating urban sprawl, fostering urban renewal, promoting sustainable real estate development, and increasing inner-city socio-economic and racial diversity.

This study has implications even where various forms of school choice already exist. For example, Milwaukee’s voucher program excludes students from families with incomes above 175% of the federal poverty level – $37,439 for a family of four in 2008-09. The threshold is apparently intended to focus resources on students from poor families. However, an unintended consequence of restrictive eligibility may be to further concentrate poor families in the inner city, while middle-class families relocate to the suburbs. When one considers the greater environmental impact of the voucher policy, a better design might allow wealthier families, including suburban families, to enroll. Subsequent migration into the city would produce positive environmental externalities in terms of reduced sprawl, urban renewal, and increased inner-city diversity.
References:


