Using Tax Parcels to Select a Location-Based Sample: An Illustration That Examines Residents' Awareness of Sex Offenders in Neighborhoods
Sarah W. Craun and Bridget Freisthler
Eval Rev 2008; 32: 315 originally published online Mar 17, 2008; DOI: 10.1177/0193841X08316110

The online version of this article can be found at: http://erx.sagepub.com/cgi/content/abstract/32/4/315
Using Tax Parcels to Select a Location-Based Sample

An Illustration That Examines Residents’ Awareness of Sex Offenders in Neighborhoods

Sarah W. Craun
University of Tennessee

Bridget Freisthler
University of California, Los Angeles

Social science research is increasingly considering place when examining social programs and policies with a spatial component. A specific research challenge involving spatial policies is how to select a sample of individuals based on their geographic locations. This article illustrates the use of geographic information systems, tax parcels, and mail surveys to target residents in varied geographic areas. A provided example demonstrates how researchers obtained a sample of respondents living within one tenth of a mile of multiple registered sex offenders. The challenges of using tax parcels to obtain addresses for apartments and mobile home parks are also explored.

Keywords: spatial sample; GIS; tax parcels; social policy; sex offenders

The uses of both spatial data and mapping have substantially increased in social science research in the previous few years, as social science fields recognize the importance of the how place can affect individual-level behaviors and well-being. In the past, selecting decent sized samples based on spatial characteristics has typically involved significant resources and time. Through the use of geographical information systems (GIS) and tax parcels, obtaining a large, spatially based sample can be done more efficiently, allowing researchers to capture a sample based solely on their location. This article describes potential research questions that can be answered with the use of parcel data, details how to identify and use location data from tax parcels, and discusses the limitations of this approach.
Spatial data allow for the integration of the concept of space in examining social dynamics that may be place based (Goodchild et al. 2000). These social dynamics with spatial aspects can be examined using sophisticated mapping techniques. Mapping is commonly done through a computer program called GIS or with census data through American Fact Finder on the census Web site. Mapping provides many benefits such as the ability to increase awareness of public concerns, such as the location of crimes, or visualize the location of services to ensure equal access across neighborhoods (Hillier 2007; Ratcliff 2002). Moreover, the creation of maps, along with appropriate spatially sensitive analytic techniques, such as spatial analysis with hierarchical models, can allow researchers to examine the interaction among individual-level problems, neighborhood characteristics, and spatial distributions (Chaix et al. 2006).

As social problems tend to occur not in a vacuum but in some location, place-based data allow for a more inclusive research model, which is likely to be more representative of the surrounding environment. Yet research studies involving place-based data and social problems tend to rely on secondary data from administrative records. This approach is useful because these databases tend to contain all reported occurrences of an event (e.g., traffic crashes, child abuse referrals, suicides), the locations of where the events occurred (e.g., street address), and minimal demographic information (e.g., age, race/ethnicity). Additionally, the databases tend to be updated on an ongoing basis for other purposes (e.g., by police to report crime statistics). However, the limited amount of information available tends to restrict the types research questions that can be answered. Spatially sampling individuals to obtain data that are not easily captured in administrative records, such as data about individuals’ opinions or perceptions, is not as common. Having a sampling plan that would not be cost-prohibitive or overly time-consuming would allow for answering a variety of research questions that currently require significant resources. These research questions might include, what do neighborhood residents think about policies that have spatial characteristics? Are residents knowledgeable about and use neighborhood-based programs? And do residents know of the impact of a potential nearby environmental hazard? These types of questions require a sampling plan that would have access to a population of addresses from which to draw a sample.

Spatial sampling is not a new concept. Geographers have devised several means of sampling over a specified geographic area for various purposes. For example, adaptive sampling procedures allow researchers to choose a sample based on a variable of interest (Thompson 2002). When it comes to
examining spatial attributes, this means that a sampling design that maximizes the chances of finding the spatial variables of interest (e.g., pollution, migrating tendencies of birds) may be chosen. This allows researchers to be more efficient in their sampling procedures as it capitalizes on known information about the location where these events or behaviors occur. In studying neighborhoods, a common approach has been to stratify neighborhoods on the primary variable of interest (e.g., poverty, alcohol outlet density; Sastry et al. 2006; Treno et al. 2008) to understand how these factors may affect behavior. In these studies, individuals have then been randomly selected either through phone or through in-person surveys. However, in the current study, where the focus was on determining how to sample individuals who lived near a person with a particular characteristic (i.e., sex offender), neither adaptive sampling nor stratification were possible, and we had to seek out an alternative sampling procedure.

The importance of spatial sampling in evaluative research can be seen when considering sex offender registration, or the policy commonly known as Megan’s Law. Megan’s Law was formulated in response to a highly publicized case where a child, Megan Kanka, was abused and murdered by her neighbor, a convicted sexual offender. The parents argued that if they had known about the neighbor’s past, then they would have been able to better protect their child. Megan’s Law augmented previous legislation by requiring the release of sex offender information to the general public to protect the public from additional harm (Office of the Attorney General 1999). To evaluate if this policy is working as intended to inform neighbors of the presence of a local sex offender, spatial sampling is crucial to select a relevant group of participants. If a diverse group of neighbors is sampled from distances around multiple sex offenders, then research can be conducted to determine if neighbors are aware of the sex offender and if protective measures are being taken.

As the spatial dynamics of social problems are recognized, policies designed to address these problems also have spatial aspects. Two prominent examples are sex offender residency restrictions and regulating alcohol outlets around schools and parks. Specifically, several states have adopted policies, which have withstood legal challenges, that restrict the locations where registered sex offenders can live (Levenson and D’Amora 2007; Zandbergen and Hart 2006). This type of spatially based policy leads to research questions where obtaining the correct geographic population is crucial. For example, are those who live within restricted areas for registered sex offenders more apt to let their children play outdoors as compared to those who live beyond the restricted areas? Are those within the restricted
areas aware that this residency restriction exists? In an example at the local level, Contra Costa County in California has passed an ordinance (Alcoholic Beverage Sales Commercial Activities 2002) that prohibits the placement of any new alcohol outlet within 400 feet of schools, parks or recreation areas, churches, alcohol or drug treatment programs, and county social service offices. Similarly, the county does not allow any new alcohol outlets to open in those crime districts with higher-than-average crime rates. For researchers, relevant questions about this type of policy include, what is the correct “distance” that outlets should be from children’s play spaces to reduce alcohol-related problems? Should different types of alcohol outlets (e.g., bars vs. liquor stores) have separate regulatory distances?

Not only social policies but increasingly social programs are considering the impact of geography on provided services. It is now commonly accepted that spatial distribution, or geography, can affect both the social interactions and social services in and between communities (Coulton 2005). Furthermore, it is known that social programs or policies may be more or less effective depending on their geographic components (Coulton 2005). Research studies using spatial data allow professionals in fields such as social work and law enforcement to target certain geographic regions for better service implementation. For example, Chahine, van Straaten, and Williams-Isom (2005) found a disconnect between neighborhood residents’ perceptions of what community-based services were available in their neighborhoods, whereas the same community organizations complained they were being underutilized. Wong and Hillier (2001) used GIS to study the location of the residences of participants of a homelessness prevention program. They determined there was a relationship between the location of the residences of those who participated in the program and the program office locations. The researchers found that different program offices had differing levels of distance traveled by the participants and that a strategically placed new program office could reduce the distance traveled by a portion of participants (Wong and Hillier 2001). Wong and Hillier (2001) utilized preexisting addresses from current participants and previous shelter residents. If the researchers expanded their research questions to ask about perceptions of the program in the general community, not the location of participants who were utilizing the program, preexisting addresses from agency files would not provide an adequate answer.

The foster care system is also attempting to develop a deeper understanding of why some neighborhoods have higher numbers of child abuse reports, and, consequently, they want to create interventions that account for this spatial disparity. For example, Freisthler (2004) examined the spatial
relationship between the rates of alcohol outlets and neighborhood child maltreatment to find that neighborhood poverty, unemployment, female-headed households, and high densities of bars were found in census tracts with high child maltreatment. In addition, there has been a developing idea in foster care that keeping foster children in the same neighborhood where they were prior to entering foster care can ease the transition into placement for children (Chahine, van Straaten, and Williams-Isom 2005; Duerr Berrick 2006). Although this practice has been debated in the literature, it is agreed on that research on the utility of this new spatially based program is unknown and that further research is needed (Crampton 2007; Duerr Berrick 2006).

Obtaining a sample based on geography can be important for environmental justice studies as well. It has been argued that most studies examining risk of populations living near environmentally hazardous materials focus on larger units of analysis, such as census tracts (Bevc, Marshall, and Picou 2007). For example, higher percentages of minorities are found in neighborhoods surrounding hazardous waste treatment facilities (Mohai and Saha 2006). When researchers aim to examine individuals rather than communities, the workload increases. In one study, when researchers obtained individual-level survey data of respondents living close to a hazardous waste site, they walked door-to-door to solicit residents (Bevc, Marshall, and Picou 2007). This method requires a large amount of time and resources to complete while limiting the ability to include multiple sites. If multiple waste sites across various geographies were examined, the amount of time and resources needed would substantially increase.

The previous literature indicates that a method is needed where a spatially based sample can be chosen with relative ease across multiple sites. In this article, we present a method that explains how tax parcels can be combined with mapping to obtain a sample of participants based on their geographic locations. Utilizing GIS to obtain a sample of potential respondent addresses from tax parcels, in which residential zoning can be determined, allows researchers to economically acquire a spatially based sample.

The Benefits of Using Tax Parcels and GIS for Social Research

Much of the previous research that has used GIS has been conducted by using some geographic unit denoted by administrative boundaries, such as a census tract or a zip code. For example, Miles-Doan (1998) used census
tracts and police crime records to determine if domestic violence varied by neighborhood, whereas Smith, Frazee, and Davison (2000) examined the relationship between face blocks (blocks on opposing sides of a street) and robbery. Researchers have also used the combination of zip codes and telephone exchanges to survey gay men about HIV and to compare characteristics of men who have sex with men living in “gay ghettos” to those who did not (Catania et al. 2001; Mills et al. 2001). If the goal is to obtain a targeted sample to determine outcomes of a location-based intervention or program or other general knowledge that is spatially based, tax parcels would be another appropriate option. For example, it was found that urban areas are more likely to have same-sex cohabitating couples than rural areas, and within those urban areas distinct clusters can be identified where same-sex cohabitating couples outnumber opposite-sex cohabitating couples (Hughes and Saxton 2006). Having a sampling method that utilizes tax parcels allows for easier targeting of respondents in areas with high and low concentrations on the variable of interest—same-sex cohabitation in this case.

A rarely employed method of obtaining a spatial sample of residents is through the use of tax parcel data. In a previous study, Lee, Moudon, and Courbois (2006) used a spatial sampling approach with parcels in their examination of environmental correlates of walking and biking. In this study, the authors had spatial requirements that limited the sample to parcels that were contained in areas that had 10 or more dwellings per acre of land and were close to retail outlets. After selecting the parcels, the researchers used telephone surveys to contact the participants. The authors argued that this type of sampling allowed for the testing of individual environmental variables while still considering the neighborhood type. However, the researchers did note that when studies used fieldwork or in-person interviews, the time involved may increase as locations could be dispersed across significant distances. Our work augments the method presented by Lee, Moudon, and Courbois by illustrating how spatial sampling using tax parcels and mail surveys can provide needed information for social science research while still reaching diverse geographic locations. Specifically, we provide the method used to evaluate a spatially based concern, awareness of registered sex offenders in neighborhoods. To answer the research question determining the percentage of residents aware of nearby neighborhood sex offenders, spatial sampling was used. This study was conducted by researchers on the West Coast on a sample taken from the population of an urban county in the Southeast United States. The method was developed to be able to remotely sample residents from many different neighborhoods with differing distances to registered sex offenders.
Methods

Background Information

The presented sampling method was developed to determine the levels of awareness in various neighborhoods about local registered sex offenders’ residences. At the time, the lack of a comprehensive and easily accessible sex offender registry in the authors’ home state required the authors to look to other locations for a possible study. Using minimal resources and wanting to obtain the necessary numbers of residents from (a) varying parts of a county and (b) varying distances to a registered sex offender was the motivation for the creation of this sampling plan.

Population of Interest

The population of interest from which we drew our sample was residents who lived in the same neighborhood as registered sex offenders. We operationalized this concept by including all parcels that were zoned residential and were within one tenth of a mile of 100 registered sex offenders. Although one tenth of a mile is much smaller than residents’ concepts of their neighborhoods (Coulton et al. 2001), the small distance was used to increase the likelihood that when respondents were asked about “the neighborhood” it would include an area where a registered sex offender lived. A comparison group was needed to determine if the awareness of sex offenders was specific to areas where sex offenders lived or if the levels of awareness were stable from neighborhood to neighborhood, which would indicate some mass perception about the pervasiveness of sex offenders in communities. Thus, the success in answering our research question would be dependent on being able to identify those residences located within certain radii of registered sex offenders.

Sampling Procedure

A total of 100 valid sample offender addresses were randomly drawn using the local sex offender registry Web site. These addresses were converted to digital maps using ArcMap 9.0 so that their addresses could be overlaid on the parcel layer. To convert the addresses, we geo-coded, or matched the sex offender addresses to digital maps of roadways, using TIGER (Topographically Integrated Geographic Encoding and Referencing system) line files, which can be obtained through the Census Bureau.
On the same map that shows the sex offender addresses, the tax parcel file was overlaid. For this study, the tax parcel information was available online to the public at the county level. Items available in the tax parcel file included the owner’s name and address, how the parcel was zoned, acreage, and sale price. These components are typical for tax files, which are generally maintained at the local level by public agencies because local governments typically manage local land use (Lee, Moudon, and Courbois 2006; Kollin et al. 1998). On this map, a radius of one tenth of a mile was drawn around each sex offender’s address (see Figure 1). All of the residential addresses within these 0.1-mile buffers were included in the pool of respondents from which the final investigational group was determined. To do this, we used a “select by location” feature in ArcMap to capture all parcels that intersected with the buffers, which had radii of one tenth of a mile. In other words, any residential parcel was selected if it intersected with the area of the buffer, which was approximately .03 sq. mi. All of the parcels that were zoned for business uses were excluded from the sample pool. From the investigational sample group pool, the researchers selected eight respondents from residential parcels within the one tenth of a mile buffer for each sample sex offender. Seven of the eight respondents per sex offender were selected at random from residential addresses. The remaining one respondent was selected at random from parcels adjacent to the sex offender’s residence. The selection of one random adjacent parcel was to ensure that the sample would include respondents who lived next door or directly behind a sex offender, with the theory that those who lived directly adjacent would have to be more invested in being aware of a convicted sex offender, and hence more likely to know of the local sex offender, than someone down the street.3

The second group from the study sample was the control group. The potential respondents in the control group must have lived at least 1 mile away from any registered sex offender, not merely the 100 sex offenders in the sample group. Figure 2 illustrates the areas from which the control group was selected. Any white area, within the county boundary, was the area from which the addresses for the control group were randomly selected. To do this, the researcher selected all parcels from within the mile buffers and then removed these from the overall parcel sample to obtain a list of possible respondents. From this list, 800 residential parcels were randomly selected to be included in the control group.

After parcels were selected, a survey addressed to “resident” of the sampled county was sent to the residential address listed on the parcel. Additionally, considering the clustering of registered sex offenders in certain neighborhoods, the sampling procedure captured some parcels that were within the one tenth of
Figure 1
Established Boundaries to Select Possible Respondents
Figure 2
Obtaining Respondents for Control Group
a mile radius of two or more registered sex offenders. A count was calculated for use as a control variable to account for this clustering.

**Special Considerations to Sampling by Parcel—Apartment Buildings and Mobile Home Parks**

Although the aforementioned section provided a method of contacting those living on a parcel that has only one single-family home, the method faces significant challenges when attempting to survey those who live in apartment buildings and mobile home parks. Criminologists have also encountered this problem when exploring crime distribution using TIGER files (Bichler and Balchak 2007). In the registered sex offender awareness study, 15% of the selected sample of registered sex offenders had an address that indicated they lived in an apartment or a room (e.g., a boarding house or hotel). This causes a problem, as neighboring apartments or mobile homes would not be picked up by the sampling procedure that looked only at parcel addresses. Therefore, mobile home parks and apartments need extra attention because the address listed on the tax parcel will not be representative of all possible addresses for the residents. For example, in mobile home parks, residents may own their mobile home while renting a space from the parcel’s owner. Therefore, multiple residents will have multiple addresses that are not indicated in the parcel tax file.

As seen in Figure 1, the sex offender in the center-left portion is located in a large parcel and has a reduced number of possible nearby parcels from which to sample. Large parcels such as these may be indicative of various possibilities. One option is that the parcel contains a single house on a large piece of land, which one would notice is distinctly different in size in comparison to the surrounding parcels. Another possibility is that the residential parcel is a mobile home park, or, third, the residential parcel could be an apartment complex.

In our study, we examined those sex offenders who, from their location on a large parcel of land as illustrated on a map, we thought could live in an apartment or a mobile home park. In addition, many sex offender addresses listed the apartment or room number, which helped indicate which parcels had multiple residences. To obtain the addresses for the multiple residences on a single tax parcel with minimal use of resources, we found it necessary to use the Internet and satellite photographs. In this study, we used Google Maps to obtain satellite views; however, any Web site that provides good- clarity satellite photos would be satisfactory. Using the zoom feature on satellite photos, one can verify if the parcel of land is a mobile home park. It is possible to see the street names by using the hybrid option on the map.
Knowing the street names of those streets nearby and the scale at the bottom of the map allowed for an estimate of those streets that were within the one tenth of a mile boundary for a more complete sample.

Once we verified that the parcel of land was a mobile home park and not just a home with a large yard, we could obtain the addresses of those nearby to the offender. The addresses of potential respondents were obtained through a “reverse telephone lookup” on the Internet. By typing the street name, city, state and zip code, a reverse telephone search produced all the addresses of residences that had a listed phone number on the specified street. Then, by using Google Maps, we could identify the location of the new addresses to determine if they were within the one tenth of a mile boundary. Using a reverse telephone lookup inherently contained some bias in that it was not possible to include those in the sample who did not have listed phone numbers. However, this biased technique was still a better option, as it produced a more complete population from which to draw our sample. The other option would have been excluding registered sex offenders who lived in mobile home parks.

As compared to mobile home parks, apartment houses provided an additional level of complexity, as the reverse telephone lookup did not produce the apartment number. For these cases, we had to make educated estimates, based on the format of the registered sex offender’s apartment number, as to what would be the other likely apartment numbers in the complex. For example, if the registered sex offender was located at Apartment 2A, one could guess that there was most likely an Apartment 1A and possibly a 3A. Yet was there a 2B or 2C? Could this numbering pattern be applied to the next parcel that also appeared to be an apartment complex? These assumptions were made in an attempt to reach the most potential respondents.

Although these estimation techniques are not ideal for contacting potential respondents in mobile home parks or apartment buildings, they can be utilized when it is not possible for a researcher to individually examine all of the addresses of interest because of geographic distances or lack of resources. However, if possible, the additional footwork of individually examining the neighboring addresses or apartment numbering schemes would provide a more comprehensive listing from which to sample.

Results—How Did Spatial Sampling With Tax Parcels Perform?

Of the 1,600 surveys mailed to addresses obtained using this technique, 11.2% \((n = 179)\) came back as undeliverable, which indicated that nearly
89.0% of the surveys were delivered by the postal service. Considering only those that were delivered, the final response rate for the mailed survey was 45.0%, with a contact rate of 49.7%. Using the method of the American Association for Public Opinion Research (2000), the contact rate was the number of returned surveys divided by the sum of the returned surveys, refusals, and nonreturned surveys (that were not listed as undeliverable).

The listed residences of 100 registered sex offenders composed the reference point for the sample selection. Using this selection technique, each offender had an average of 1.31 of his or her neighbors’ surveys (from the investigational group) returned as undeliverable. Every offender had at least one returned as undeliverable, and only one offender had all eight addresses returned as undeliverable. Further investigation of this single offender found that his listed address was a local motel in a more industrial part of town.

Of the almost 200 surveys that were returned by the postal service as undeliverable, the most frequently listed reason for return was “vacant address” (36.4%). Other frequent reasons for return included “no such street address” (22.2%), “not deliverable as addressed” (15.0%), and “attempted, not known” (12.7%). There was a difference in the percentage of surveys returned as undeliverable between the two groups, with the investigational group having 16.4% of its sample returned, whereas only 6.0% of the control group’s surveys were undeliverable, $\chi^2(1, n = 1,600) = 43.34, p < .001$. Table 1 illustrates the distribution of returned surveys by group, where there was a significant difference by group for only two reasons—“attempted, not known,” $\chi^2(1, n = 179) = 9.19, p < .01$, and “not deliverable as addressed,” $\chi^2(1, n = 179) = 5.80, p < .05$. In addition, 13.0%

<table>
<thead>
<tr>
<th>Reasons for Return by Postal Service (%)</th>
<th>Control Group ($n = 48$)</th>
<th>Investigational Group ($n = 131$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attempted not known**</td>
<td>0.0</td>
<td>16.8</td>
</tr>
<tr>
<td>Insufficient address</td>
<td>4.2</td>
<td>4.6</td>
</tr>
<tr>
<td>No mail receptacle</td>
<td>8.3</td>
<td>5.3</td>
</tr>
<tr>
<td>No such number</td>
<td>29.2</td>
<td>19.1</td>
</tr>
<tr>
<td>Not deliverable as addressed*</td>
<td>25.0</td>
<td>10.7</td>
</tr>
<tr>
<td>Vacant</td>
<td>31.3</td>
<td>36.6</td>
</tr>
<tr>
<td>Other or no reason recorded</td>
<td>2.1</td>
<td>6.9</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.
(n = 24) of the returned surveys were addressed to apartments, and all of the 24 surveys were part of the investigational group.

To further investigate this, we examined the neighborhood characteristics of the individual addresses, measured by Census Summary File 3 data at the census tract level using multiple one-way ANOVAs with a Bonferroni adjustment. The surveys that came back as undeliverable were more likely to come from addresses in neighborhoods with higher percentages of African Americans, higher percentages of poverty, higher numbers of people per square mile, higher percentages of single parents, and lower median incomes (p < .001) than the addresses where surveys were not

Table 2
Neighborhood Characteristics of Sampled Addresses by Response Group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Responding Addresses (n = 631)</th>
<th>Nonresponding Addresses (n = 790)</th>
<th>Undeliverable Addresses (n = 179)</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>White (%)*</td>
<td>72.2</td>
<td>70.4</td>
<td>64.2</td>
<td>9.24</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>African American (%)*</td>
<td>21.0</td>
<td>22.6</td>
<td>28.1</td>
<td>7.53</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Asian (%)</td>
<td>2.7</td>
<td>2.5</td>
<td>2.3</td>
<td>1.62</td>
<td>.20</td>
</tr>
<tr>
<td>Native American (%)</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.15</td>
<td>.86</td>
</tr>
<tr>
<td>Other race (%)*</td>
<td>2.3</td>
<td>2.5</td>
<td>3.4</td>
<td>12.90</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Two or more races (%)**</td>
<td>1.7</td>
<td>1.6</td>
<td>1.5</td>
<td>3.97</td>
<td>.02</td>
</tr>
<tr>
<td>Hispanic (%)*</td>
<td>4.8</td>
<td>5.2</td>
<td>6.2</td>
<td>6.40</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Households in poverty (%)*</td>
<td>7.7</td>
<td>8.2</td>
<td>12.8</td>
<td>22.37</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Median income (dollars)*</td>
<td>60,826</td>
<td>58,932</td>
<td>48,023</td>
<td>25.45</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Single parents (%)*</td>
<td>11.9</td>
<td>12.8</td>
<td>16.5</td>
<td>15.18</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Median years at residence</td>
<td>4.0</td>
<td>4.1</td>
<td>3.9</td>
<td>0.59</td>
<td>.56</td>
</tr>
<tr>
<td>Lived at same residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years ago (%)</td>
<td>42.0</td>
<td>42.6</td>
<td>41.2</td>
<td>1.26</td>
<td>.28</td>
</tr>
<tr>
<td>Average density (people per square mile)*</td>
<td>1,382</td>
<td>1,375</td>
<td>2,056</td>
<td>11.72</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Vacant units (%)</td>
<td>19.3</td>
<td>19.1</td>
<td>18.6</td>
<td>0.27</td>
<td>.76</td>
</tr>
<tr>
<td>Foreign born (%)</td>
<td>8.2</td>
<td>8.3</td>
<td>8.7</td>
<td>0.61</td>
<td>.54</td>
</tr>
<tr>
<td>On public assistance (%)*</td>
<td>1.8</td>
<td>2.1</td>
<td>2.8</td>
<td>8.91</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Owner-occupied housing (%)*</td>
<td>72.7</td>
<td>71.7</td>
<td>62.1</td>
<td>18.07</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Elderly (%)</td>
<td>6.4</td>
<td>6.6</td>
<td>6.2</td>
<td>1.45</td>
<td>.23</td>
</tr>
<tr>
<td>Male (%)*</td>
<td>49.6</td>
<td>49.5</td>
<td>50.3</td>
<td>5.03</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Unemployed (%)*</td>
<td>3.8</td>
<td>4.1</td>
<td>6.7</td>
<td>14.34</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*Undeliverable addresses are significantly different than both respondents and nonrespondents. **Undeliverable addresses are significantly different from only respondents.
returned as undeliverable (see Table 2). In other words, this method works less well to sample those residents who may already be underrepresented in social science research. A word of caution is needed in interpreting the results of Table 2. It is possible that we are creating a fallacy in attributing census tract data to individuals. Although it would be more appropriate to examine differences at the census tract level among responders, nonrespondents, and undeliverable addresses, this is not possible in that we have no census tracts that contained only undeliverable addresses.

Discussion

As technology advances and social science research takes advantage of this technology, there is an increased ability to answer complex research questions. For example, computer-based mapping has allowed researchers and practitioners alike to view the spatial distribution of social concerns, such as suicide or the clustering of fast food restaurants (Baller and Richardson 2002; Block, Scribner, and DeSalvo 2004). Here, we show that technology advancements can also be employed to obtain samples to answer research questions that are based in respondents’ personal perceptions. In this article, we have made an argument that a cost-effective technique is needed to obtain samples of individuals based on their geographic location. We then explained how GIS can be utilized in combination with tax parcel files to obtain such a sample. Finally, we provided an illustration of how this method has worked for a research project examining residents’ awareness of registered sex offenders in different neighborhoods.

Using the technique of combining GIS and tax parcels, we were able to obtain the addresses of 1,600 potential respondents from various areas of a single county in a state distant from the authors. The total number of undeliverable addresses was only 11% of the initial mailing. Beyond determining only the number of surveys that were returned, we further explored if there were differences in why the surveys were returned, according to the U.S. Postal Service. There were two reasons for the return of surveys where the distributions significantly differed. The reasons provided were delivery was “attempted, not known,” and the survey was “not deliverable as addressed.” It is interesting that the control group had no surveys returned for the reason of “attempted, not known.” The U.S. Postal Service (2007) explains this code by saying the addressee was not known at the address. We are unsure why there would be a difference between the groups for this reason, as all 1,600 mailings were addressed to the “resident” of the sampled county. We
surmise it may have been differing processing by differing post offices. Furthermore, “not deliverable as addressed” is explained in the Postal Service manual as mail that is undeliverable at the address given, mail for residents for whom there is no change of address on file, or mail for residents for whom the expiration has passed on the forwarding order. Again, we are unsure of why these differences occurred. Future work may be better able to explain these differences or determine if they were isolated to our study alone.

In an examination of undeliverable surveys by study group, there were significant differences in that those surveys that went to the investigational group were more likely to be returned as undeliverable. In addition, it was more difficult to contact potential respondents in areas where there is more concentrated disadvantage. For example, surveys that came back as undeliverable were typically from census tracts that had more people per square mile and higher percentages of households in poverty than did the census tracts of responders and nonresponders. In addition, those in apartments are more difficult to establish contact with using this method, which is supported by the statistically significant lower percentage of owner-occupied housing in the census tracts attributed to undeliverable addresses.

Although this sampling method may have performed slightly worse in poorer areas, the understanding of the result can lead to better planning in the implementation of this technique. For example, during the planning stage of a research project, an awareness that additional time and resources may be needed to contact those who live in poorer areas can be utilized in asking for additional resources. Conversely, researchers may want to utilize this knowledge by oversampling in poorer areas as well.

To our knowledge, this is the first instance of researchers concurrently using GIS and tax parcels to obtain a sample for a mail survey. Clearly, this method can benefit from further refinement for a smoother implementation on a larger scale. One proposal for improvement stems from the work of Zandbergen and Hart (2006), who used parcel files to determine the percentage of housing available for registered sex offenders in a county where sex offender residency restrictions may be implemented. In their work, although they did not contact parcel residents, they determined if a parcel was suitable for occupancy by ensuring that the parcel file listed the year built and posted the square footage of the living area and that the building value was assessed. Considering all three data points was not done in the current study, as only zoning was considered from the tax parcel file, along with the address. Perhaps, if the current study had utilized all three data points, the number of undeliverable surveys would have been reduced.
Limitations

Although we contend that this method can be quite beneficial to program evaluation and policy research, we note some of the limitations in our implementation. As with all secondary data sources, researchers using this method must rely on others for the parcel data to be correct. Human error in data entry for the parcels can influence the number of contacts that can be made with potential respondents. In fact, Zandbergen and Hart (2006) found in a sample of 200 parcels that were deemed to be occupied from the data alone that 3% were either vacant or under construction. In the current study, we had 11% of the mailed surveys come back as undeliverable. If the researcher is reasonably close to the geographic area being studied, then inspection of the parcels that come back as undeliverable is possible, albeit time-consuming. However, if the numbers are too great or the distance is too far, this limitation can prove difficult. A pilot test can provide an estimate of the number of undeliverable questionnaires in a study area, which can then be factored into the proposed sample size. Moreover, as mail surveys struggle with response rates, future studies utilizing this method may want to consider attempting a telephone contact to follow-up.

An additional limitation focuses on the changing parcel files. Depending on the area of study, new developments may be outpacing the ability to enter data into the tax parcel file. This would hinder the ability of the researchers to obtain a sample representative of their study population. Tax parcel files should indicate the date of the last update for a relatively easy check of the time of the data entry. Conversely, if the tax parcel files are updated too quickly, they may record the address of a newly zoned residential parcel where construction is not yet completed. Therefore, a residential address would be attributed to a residence that is not yet suitable for habitation.

Finally, there was a limitation in specifying to whom the survey was directed. In the current study, the survey was directed only to “resident,” and therefore the authors had no control over which household member answered the survey. The coding provided by the U.S. Postal Service on the undeliverable surveys suggests that directing the survey to “resident” may have increased our number of undeliverable surveys. This county’s particular parcel data included the name and address of the owner of the parcel. If needed, the survey could have been addressed to the name of the owner, if the owner lived at that parcel. However, it was not critical to this survey to have the owner complete the survey; rather, our focus was on contacting a resident of the parcel at the time of the mailing. If a name of the current resident is needed for a research project, a reverse telephone lookup may be used to obtain the name of the person listed in the telephone book.
Conclusion

The implementation of spatially based policies and programs, and hence the need to evaluate such policies and programs, necessitated the development of a method to efficiently sample respondents based on their geography. In this article, we have illustrated a method to be used when the research question is spatially based and requires the responses of individuals. Using administrative data, namely, tax parcels, census data, and other secondary data sources, researchers have the ability to contact potential respondents to answer a variety of research questions, including such topics as the impact of space-based policies or an evaluation of a local organization’s attempt to reach the neighboring community. This allows for a more nuanced approach to evaluation that shows not just whether or not a certain policy or program is effective but also whether or not that same policy or program is more effective in certain locations than in others. While acknowledging the limitation regarding reaching a smaller proportion of respondents in poorer areas and in apartments and mobile homes, we believe that with future implementations, this method can be further refined to better serve the needs of researchers.

Notes

1. Readers can contact the first author for a copy of the research study, which led to this sampling design.
2. Only registered sex offenders whose address had a date of verification by law enforcement were included in the sample.
3. The splitting of the sampling into 7 and 1 would have required special weighting; however, in future analyses it was found that the adjacent variable was not relevant and therefore it was not used in any multivariate analyses.
4. To increase the response rate for the mail survey, we utilized the suggestions of Dillman (2000). In this study, we sent an introductory letter, an initial survey a week later, a follow-up postcard, and then a second wave of surveys. Two dollars were included in every initial survey, and those who returned completed surveys were entered into a drawing for a cash prize.

References


Sarah W. Craun is an assistant professor in the College of Social Work at the University of Tennessee.

Bridget Freisthler is an assistant professor in social work in the School of Public Affairs at the University of California, Los Angeles.