Prospecting for (Campaign) Gold

Wendy K. Tam Cho University of Illinois at Urbana-Champaign
James G. Gimpel University of Maryland

Campaigns and political parties are faced with the immensely important practical challenge of financing their efforts. Raising money is instrumental to all other aims. In recent years, this task has been complicated by the need to enlist ever greater numbers of contributors to raise ever larger sums of money. At the same time, fundraising burdens are eased a bit because contributors flock together. That is, campaign contributing is a spatially dependent phenomenon, associated with affluence and the presence of networks. Accordingly, geospatial tools provide a helpful method for understanding and predicting where contributions can be most successfully mined.

Few political campaigns in America can be waged without monetary donations. Though one cannot discount the value of a high-quality candidate, the quality of a candidate is largely defined by the quantity of resources he or she attracts. Not that every well-financed candidate wins, of course, as moneyed candidates go down to defeat in every election cycle (Steen 2006). Nonetheless, few would question the wisdom of establishing a well-oiled fundraising operation early in the campaign cycle. The returns from such an investment are undisputable. Indeed, even after a successful election, seeking campaign donations is still imperative to retire debt. The chase for money never really stops (Herrnson 2006), prompting one former U.S. Senator to editorialize in frustration:

> In 2004, my last year in the Senate, we had Thursday policy lunches at which experts on both sides of a hot topic would make short presentations and we would hammer out policy. But from the beginning of the summer of that year until that fall's elections, policy lunches were canceled so that senators could go to their parties' headquarters and call all over the country, begging for money.
>
> The result of this nonsense is that almost one-third of a senator's time is spent fundraising. The Senate schedule calls for morning-long committee hearings (which many times extend into the afternoon). In the afternoon, there is debate on the floor and votes. Little time is left for talking to the staff or to constituents, answering mail, phone calls, etc. Every evening there is an average of three receptions or fundraisers, followed by three breakfasts or fundraisers the next morning. (Hollings 2006)

Despite the unremitting need to fundraise, how to successfully prospect for campaign gold is far from simple or intuitive as evidenced by the spirited but fruitless efforts of many would-be officeholders. Experienced fundraisers who possess impressive know-how charge a high premium for their services, and tapping into this knowledge is often limited to incumbents and high-profile candidates. Fundraising aptitude is cultivated and learned, not inborn (Krasno, Green, and Cowden 1994; Squire and Wright 1990). The urgent need to learn these skills is vastly magnified by campaign costs that are steadily on the rise, coupled with existing donation limits that make it essential for political elites to raise money from an ever larger number of donors (Jones and Hopkins 1985). As Senator Hollings has noted, this intense pressure means that candidates must allocate increasingly large amounts of time and resources to asking for money, leading to complaints that fundraising pushes out more worthwhile activities such as discussing the issues, meeting with ordinary voters, and

Wendy K. Tam Cho is associate professor of political science and statistics and senior research scientist at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, 361 Lincoln Hall, 702 S. Wright Street, Urbana, IL 61801 (wendycho@uiuc.edu). James G. Gimpel is professor of political science, Department of Government. University of Maryland, 3140 Tydings Hall, College Park, MD 20742 (jgimpel@gvpt.umd.edu).

Thanks to Luc Anselin, Steve Ansolabehere, Pierre Goovaerts, Carol Gotway, Don Green, Laura Hussey, John McNulty, and Lance Waller for helpful comments.

American Journal of Political Science, Vol. 51, No. 2, April 2007, Pp. 255–268

 ISSN 0092-5853

legislating. Clearly, both disadvantaged challengers and overwhelmed incumbents would benefit if the burden of prospecting were eased. One path to promoting a more citizen-based electoral system and thus broadening a vital aspect of participatory democracy beyond traditional elites might be to even out the information playing field that guides would-be campaign prospectors.

Exploring the role of geography may prove to be especially propitious in this endeavor (Cho 2003; Gimpel, Lee, and Kaminski 2006). After all, the important role of location and spatial diffusion in grassroots mobilization has long been recognized (Huckfeldt and Sprague 1992). In mobilization efforts, voters are contacted, and they in turn influence their compatriots (e.g., other members of the household and neighbors). A similar dynamic is at play in the campaign finance realm (Cho 2003). The challenge for a campaign is to implement tactics that exploit the social connections and informal contacts of supporters (Huckfeldt and Sprague 1992, 84). These connections are geographically constrained, since acquaintanceship and social influence tend to weaken with distance. In fact, many political phenomena have been linked to social dynamics (see, e.g., O'Loughlin and Anselin 1996; O'Loughlin, Flint, and Anselin 1994; Shin and Agnew 2002; Starr and Most 1983).

Precisely *how* geographic understandings play into strategies of fundraising is a rather different question. Across the landscape of U.S. elections, votes are far more plentiful than dollars. Donors are so few in number that the average neighborhood contains almost none. Members of the same household are known to contribute to the same campaigns, but efficient strategies demand the presence of nearby connections beyond the immediate proximity of family members. The task of prospecting for campaign gold is further complicated by the fact that the spatial distribution of campaign donors and the distribution of electoral support are far from isomorphic (Gimpel, Lee, and Kaminski 2006; Herrnson 2006). Because the vast majority of voters give no money at all, looking to the locations that are most lopsidedly "red" or lopsidedly "blue" is minimally helpful in deciding where to "rationally prospect" for campaign funds.[1] Moreover, while campaigns have relied upon direct mail solicitation for decades, to date, far more money is raised at events, on a face-to-face basis, than through direct mail contact (Ansolabehere, deFigueiredo, and Snyder 2003; Jones and Hopkins 1985).

Detecting buried campaign treasure is closely tied to the residential locations of individual contributors and the social relationships among them. To be sure, neighbors tend to resemble one another in both socioeconomic and demographic attributes because similarly situated individuals tend to cluster. A related pattern is that many campaign donors are located close to one another, residing in the same neighborhoods or closely adjacent neighborhoods. And political organizations find it efficient to prospect for new donors in the same locations where the established contributors are located. This pattern is reinforced by campaigns that have been slow to develop significant contribution sources outside traditional centers of contributor activity (Verba, Schlozman, and Brady 1995, 158). Presently, candidates return repeatedly to locations where a culture of contributing has developed and participation through campaign giving has become a tradition.

The tendency for giving to be habitual generates a geographic pattern of financial support, greatly enhanced by the penchant to hold fundraising events at the same locations, and even the same households. More remote areas are unattractive not only because of their smaller potential donor pool, but also due to burdensome travel costs and weaker, or nonexistent, traditions of giving.[2] The "stickiness" of location is analogous to the kind of agglomeration forces that concentrate economic activity in certain places. Once a neighborhood becomes a center of contribution activity, donation history becomes as important as current conditions in predicting where prospecting will succeed, a fact recognized by professional fundraisers (Gerber, Green, and Shachar 2003; Green and Shachar 2000; Smith 2005; Ware 1992.)[3] Perhaps this is why an important recent survey of small donors has shown that their economic and demographic profile is not significantly different from larger donors (Institute for Politics and the Internet, 2006).

Here, we aim to demonstrate the enormous value of considering the spatial dimensions of campaign contributions in understanding and yielding predictions.[4] We begin by presenting a series of hypotheses about how

[1] "Blue" and "red" here refer to the Democrats and Republicans, respectively. This color-coding scheme appears to have come into vogue after the 2000 presidential election. Media outlets have taken advantage of color since color television became common. However, prior to 2000, there was no consensus on color schemes. In 2000, for the first time, all of the major media outlets employed the red-blue color scheme. In so doing, they popularized the term "red and blue states."

[2] We note that these campaign strategies are irrelevant to self-financed candidates, who are becoming increasingly common (McCarty, Poole, and Rosenthal 2006; Steen 2006). All the same, it is difficult to envision a time when most candidates would be self-financed.

[3] This behavior perhaps explains why, as early as the 1930s, the Republican Party implemented "continuous pledge giving" as a means of rebuilding party strength following the devastating losses of 1932 and 1936, with the idea of developing steady contribution habits among supporters (Ketchum 1938).

[4] We accept the notion that prediction is not always explanation. A variety of explanations may lie at the basis of the geographic variability in political behavior, and we will not have room to delve into those in this particular article. But we also reject the idea that "mere" prediction is of little value. Campaign managers care more about where to prospect than why they should prospect in a particular

and why geography may play a role in campaign finance. We then describe the data we use to examine our theories. The social complexities that surround contributors are at odds with "conventional" statistical models that assume the independence of observations and choose to focus on individual attributes as predictors of the propensity to give. Accordingly, we next present a method for spatial interpolation and explain how it applies to this research topic. We demonstrate how these techniques might allow a researcher to gain insight on pertinent characteristics such as neighborhood structure and the expanse of contributor culture, variables that help define the flow of campaign contributions from various locales. We then map our results and comment on the implications of our findings.

Campaign Networks

Related research has shown that campaign contributions are heavily reliant upon networks that are specifically groomed by parties and campaigns for the purposes of efficient resource extraction (Brown, Powell, and Wilcox 1995; Francia et al. 2003; Smith 2005). These networks are valued because they are geographically concentrated, allowing for the cost-effective harvesting of campaign donations. By bringing together multiple contributors to a single location, at a single event, the campaign can avoid the costly chore of having to visit every contributor singly.

Networks built around professional occupational associations in specific locales are the common target of fundraising requests, as physicians, realtors, lawyers, and farmers can be counted upon to donate to multiple candidates, across multiple elections. Ethnic networks are another important example of a ready-made contribution base that can be efficiently tapped because members of ethnic groups often reside in close proximity and tend to cluster (Cho 2003). Friendship networks are a third type that is of value, chiefly because they are the product of initiative and choice, certainly more so than a professional association or an ethnic network. These networks commonly unite members who share similar values, making them an uncomplicated basis for the organization of fundraising events. The idea of networks occupying discrete spaces is similar to the concepts underlying segregation indices (Duncan and Duncan 1955; Morrill 1991, 1995; Wong 1998, 1999, 2004).

place. The clear demonstration that there is spatial dependency in observations of campaign giving and that this spatial pattern extends over a specific range of territory are important first steps even if they do not yet identify a particular underlying causal mechanism.

Tapping into a network is clearly attractive as this maximizes the return from one's effort. Rather than concentrating on gaining a single socially isolated campaign contribution, the effect of tapping into a network may result in hundreds of enduring contributions. Network members see each other on a regular basis, and this regularity of interaction is instrumental in building trust and friendship. Networks also bring together like-minded or homophilous people who may share a commitment to particular political values or to a particular party (Lazarsfeld and Merton 1954; McPherson, Smith-Lovin, and Cook 2001). The network itself inclines individuals to give more in order to please, impress, and compete with other network members. A donation request from a close friend or business associate is more likely heeded because of the social pressures that are in play (Freeman 1997, 161). As we have noted above, campaigns are also drawn to prospect in areas with reputable networks. Often these areas are home to recognizable political entrepreneurs who are poised to host successful fundraising events.

A campaign's dependence on these social ties to build a donor pool confines prospecting activity to the places where they have built relationships. The geography of these relationships helps us to understand why political parties and candidates return time-and-again to the same locations even when it might seem beneficial to prospect elsewhere. Networks tend to reinforce the existing spatial distribution of campaign giving, and their presence makes some locations worth visiting when they might not be in the absence of a network. As one Texas campaign's finance director put it:

> A lady named MM makes Laredo worth going to. She knows everyone in the community and calls each of them personally to contribute to the event and then she drives to their offices to pick up the contributions herself. If we had a MM in every community we could raise $100,000+ in every city in Texas!
>
> This also explains why we raise so much more money in San Antonio than Dallas. It's because of GP and JS in San Antonio. Together, they raise us about $650,000 per year. We raised almost $400,000 at one event in October in San Antonio and then another $190,000 in March. In Dallas, our office works much harder at the events and we still barely make $200,000. It's all about the personal contact from a respected peer. It's easy to throw a letter away—it's hard to tell a community leader no. (Kirchmeyer 2006)

As a small and relatively poor city on the state's border, one would not expect Laredo to be a fundraising hot spot for campaigns for either party. Yet because of the presence of political entrepreneurs and their reliable networks, such inauspicious locations can be surprisingly productive. The existence of locations like Laredo demonstrate that campaign contributing is not the result of an aspatial process or atomistic actors behaving in isolation from one another. Nor is campaign giving predicted solely by mapping the spatial distribution of wealth, education, or professional employment (Gimpel, Lee, and Kaminski 2006).

The foregoing considerations from the theory and practice of campaign fundraising all suggest that we should expect a commanding geographic dimension to campaign giving. To the extent that contributions emerge as a function of the contribution activity of nearby others, perhaps the geographic distribution of contributions can be helpful in predicting where donations are most likely to emerge in the future. Because of the "pervasive geographic constitution" (Agnew 1990) underlying the solicitation and harvest of campaign donations, we suggest that campaigns can be guided by spatially predictive models that can guide the prospecting effort. Moreover, as we will see, these methods help us understand the patterns of campaign donations by providing a sense of neighborhood structure and the extent of community ties in various locales.

Data and Methods

Our data originate from the confidential fundraising files of the *Texans for Rick Perry* political campaign organization (hereafter *TRP*). Perry assumed the governorship of Texas in 2000, ran successfully for reelection in 2002, and won a second full term in 2006. In the 2006 contest, he was challenged by a Democratic nominee, as well as two well-known Independent candidates. The data we utilize are especially attractive because unlike the contributions recorded by the Federal Election Commission (FEC), the *TRP* files do not contain a cutoff for contributions below a certain amount, and there are no legal fundraising limits in Texas state campaigns. Central to our goal here, each contributor can be spatially located by his or her street address. Figure 1 shows the locations of the contributions in our data set.[5] The theoretical considerations we have sketched out suggest that these points are distributed, not in a spatially random fashion, but irregularly, primarily according to the proximity of network communicants, the income levels of the places where donors reside, and other characteristics associated with giving (e.g., population size).

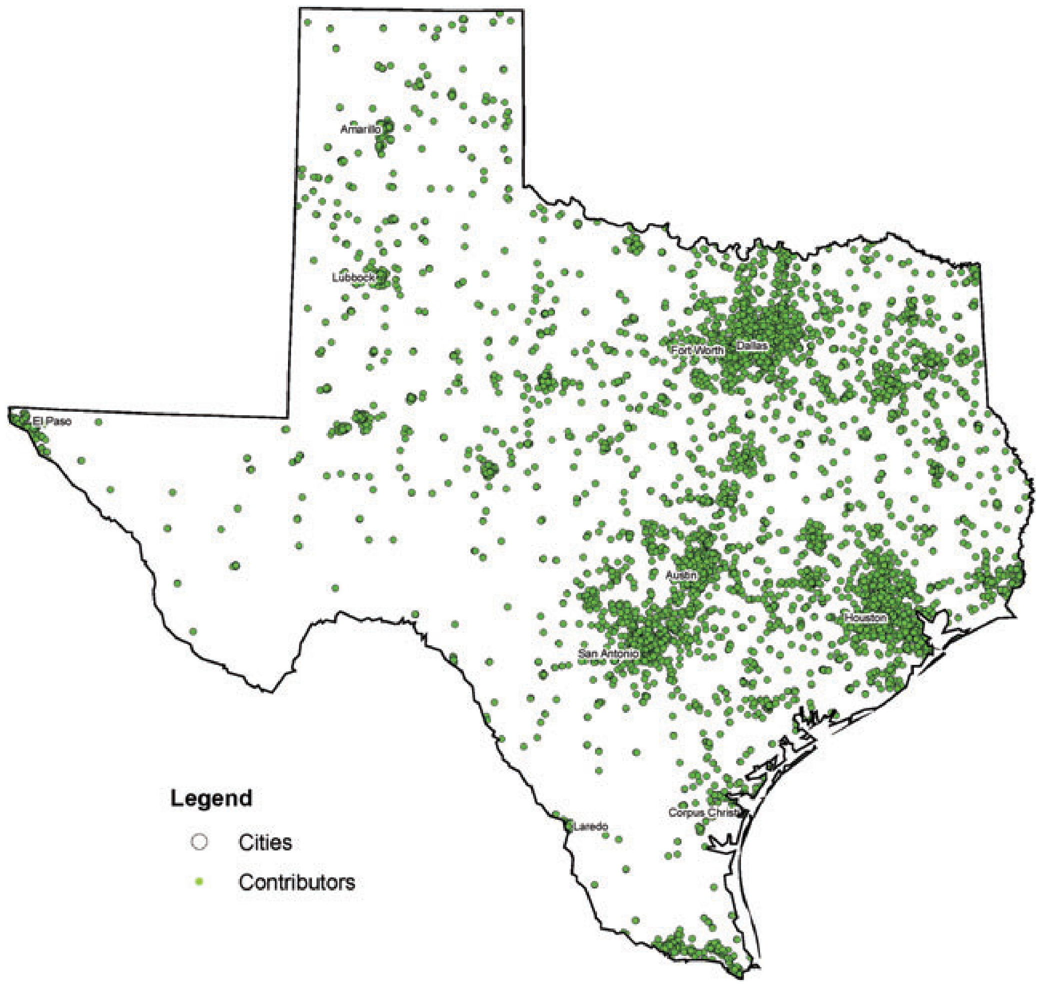
Since our data are spatial, we employ a geostatistical method to analyze them. In particular, we utilize a method called kriging to incorporate the spatial locations of our contributions into the model. We seek to use the contribution amounts and their location to estimate a "contribution surface," which we can use in turn to assign theoretical contribution amounts to locations where a contribution has not been observed as well as assess the prediction's uncertainty. The result is a "best fitting" model of the global (geographic) variation across the surface, in the sense that the predicted values exhibit the least distortion of the actual spatial structure of the observations. The central theoretical assumption is that proximate measurements are more similar than those farther apart. In this particular application, because our data represent individual cases, they have the advantage of avoiding problems associated with aggregation: chiefly the ecological inference problem, and its related modifiable areal unit problem (MAUP). Moreover, here, space is not fragmented into arbitrary areas, but is treated continuously, providing more precise information about the geographic distribution of contributions than is provided by aggregated data.

Kriging is a class of geostatistical methods that originates from earth science applications, and specifically from the work of Matheron (1963), who named it after the South African miner D. G. Krige, an engineer who pioneered the technique in the field (Cressie 1989, 1993; Stein and Corsten 1991). Kriging is now being used in other disciplines, to study animal populations, to examine disease epidemiology, environmental risks to health, housing prices, and similar georeferenced phenomena. This is the first attempt that we are aware of to apply it to any type of political behavior. What these applications have in common is that the variables of interest are regionalized, e.g., they are distributed in space (Matheron 1963). Regionalized variable theory suggests that the variation in a spatially distributed variable can be explained by three components: (1) a structural component characterized by a constant mean, or trend; (2) a random but spatially dependent component, which is the variation of the regionalized variable; and (3) spatially uncorrelated random error.

An intuitive grasp of what kriging does is displayed in Figure 2, which shows five points observed as part of a larger set of measures. The value at the sixth unmeasured point, located at (3, 3), is estimated as a weighted sum of its five measured neighboring points, where the weight

[5]The geocoding process failed to identify street-level addresses for approximately 5,000 of the 27,000 contributors. This was largely due to P.O. Boxes and Rural Routes for which a valid street range was impossible to determine. For these contributions, we assigned them to a randomly selected address in their zip code area.

FIGURE 1 Location of Contributors, 2004–2006



is determined as a function of the distance between the points and accounts for the pattern of spatial variability in the data. Kriging essentially uses a sample of points to construct a surface that can be used to predict values at unobserved locations.

Suppose that we have contribution observations $Z(s_1)$, $Z(s_2), \ldots, Z(s_N)$, where $Z(s_1)$ is the contribution amount at location $s_1 = (x_1, y_1)$, $Z(s_2)$ is the contribution amount at location $s_2 = (x_2, y_2)$, and $Z(s_i)$ is the contribution amount at location $s_i = (x_i, y_i)$, where (x_i, y_i) locates the ith observation as a point in a plane. These locations are distributed irregularly throughout our domain of interest, D. Here, the domain is the state of Texas. We can say, then, that our data are a partial realization of a random process or random field,

$$\{Z(s) : s \in D\}, \qquad (1)$$

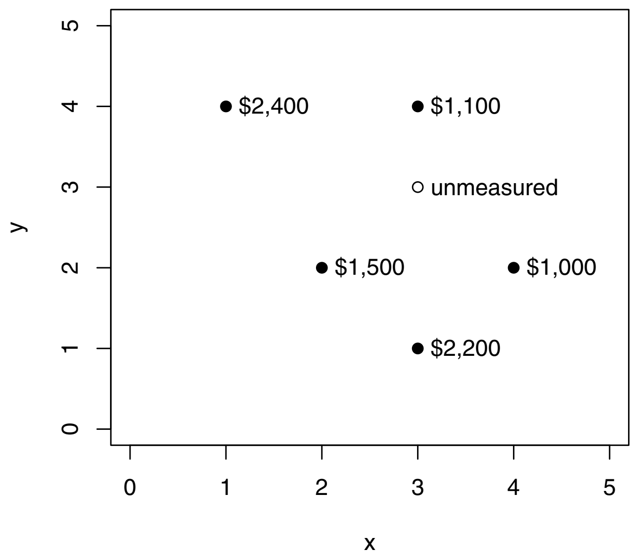
where s varies continuously throughout D. The observations are contributions at a set of locations, $Z = [Z(s_1), \ldots, Z(s_N)]'$. We can use these observations to construct a contribution surface to predict the value of the $Z(\cdot)$ process at some unobserved location, $Z(s_0)$, $s_0 \in D$.

If proximate observations are more similar than nonproximate observations (i.e., the data exhibit spatial dependency), then we expect there to be increasing variation in the pairwise differences in contributions as the distance between those pairs increases. In geostatistical applications, this spatial autocorrelation is quantified through a function, $\gamma(\cdot)$, called the *semivariogram*, which models the pattern of spatial variability.[6] The semivariogram is one half of the variogram function, $2\gamma(\cdot)$, which is defined by the equation

$$\mathrm{Var}(Z(s_i) - Z(s_j)) = 2\gamma(s_i - s_j),\ s_i, s_j \in D. \qquad (2)$$

[6]The existence of spatial autocorrelation is more formally defined by the moment condition, $\mathrm{Cov}(y_i,y_j) = (y_i,y_j) - \mathrm{E}(y_i,y_j) \neq 0$, where y_i and y_j are observations on a random variable at locations i and j in space.

FIGURE 2 **Example of Sample Locations and an Unmeasured Location on a Planar Surface**



We may estimate (2) by taking the average of all the observed squared differences, $[Z(s_i) - Z(s_j)]^2$, for pairs of observations that are the same distance apart. This gives rise to the classical semivariogram estimator,

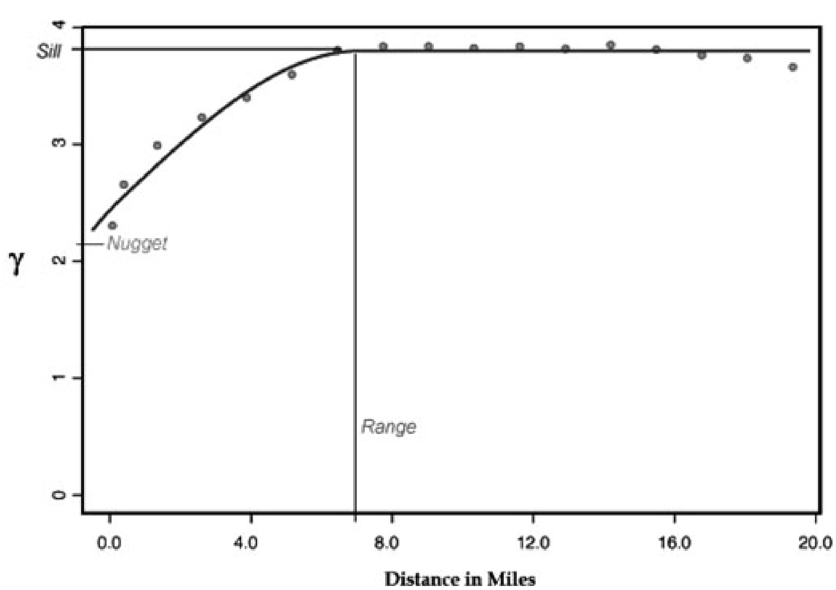
$$\hat{\gamma}(h) = \frac{1}{2 \mid N(h) \mid} \sum_{N(h)} (Z(s_i) - Z(s_j))^2, \quad (3)$$

where $N(h)$ is the set of distinct pairs separated by distance h,

$$N(h) = \{(s_i, s_j) : s_i - s_j = h\},\ i,\ j = 1, 2, \ldots N, (4)$$

and $|N(h)|$ is the cardinality of $N(h)$. For irregularly spaced data such as the set we are considering, one might also modify this definition to $\{(s_i, s_j): s_i - s_j \in T(h)\}$, where $T(h)$ is a tolerance region rather than an exact distance value. Figure 3 shows the empirical semivariogram for our data. The value of the semivariogram function is graphed against the Euclidean separation distance between observations. Typically, one specifies the *lag spacing*, or the distances at which the semivariogram is estimated, and the total number of lags or bins where the semivariogram will be estimated. Specification of the tolerance regions and the number of bins is flexible. The size of the tolerance region is chosen so that it is small enough to retain sufficient spatial resolution to define the structure of the semivariogram while also large enough so that each bin retains a sufficient number of paired differences to ensure that the empirical semivariogram at each point is well estimated (Journel and Huijbregts 1978). Generally, in practice, the maximum lag distance is approximately half of the maximum separation distance. At larger distances, the semivariogram estimates rely on points at the ends of the domain, which usually encompasses very few paired distances and wide variation in the estimates (Journel and Huijbregts 1978; Waller and Gotway 2004). In our estimation, we employed a *lag* of 1 mile with 15 total lags. That is, each bin includes a one-mile range of

FIGURE 3 **Spherical Model Fit to Empirical Semivariogram of Campaign Contributions**



distance and the plotted value is an average of the values that fall into that bin.

From this plot, we can glean information about the continuity and variability of the spatial process. We expect positive spatial autocorrelation to exist at closer distances. At some point, as the distance between observations increases, we expect the semivariogram to level off. This point, termed the *sill*, represents the maximum variance in the observed data. The distance where this maximum is reached, and the distance at which the spatial dependency disappears, is called the *range*. Distances farther than the range apart are spatially uncorrelated as reflected by the near constant variance between pairs shown in the semivariogram. The range also sets the limiting distance within which interpolation via kriging is worthwhile (Webster 1985, 19). The *nugget* is the point at which the model intercepts the y-axis. While theoretically the nugget should always be zero because this represents the semivariance value when there is no distance between points, nuggets greater than zero can result as a consequence of measurement error, misspecification of the semivariogram model, or due to variation at spatial scales smaller than the sampling distances of the data (Cressie 1993; McBratney and Webster 1986). Examining the semivariogram is a first step in modeling as we use it to determine the extent of autocorrelation.

Next, we parametrically model the empirical semivariogram. The parameters from this semivariogram model are then used to predict values for unsampled locations. In Figure 3, the points represent values calculated from our data set and are used to estimate the unknown semivariogram. The curved line is a theoretical semivariogram that is fit to our data, so that the function closely reflects our empirical semivariogram. The theoretical semivariogram is from a specific parametric family. In this case, it is the set of classical or spherical semivariogram functions.[7] Once the spherical semivariogram family is chosen we find a vector of parameters that produces a close fit to our data. As we can see from the plot, the theoretical semivariogram (range = 7.4 miles, sill = 3.84, and nugget = 2.83) produces quite a good fit to the empirical semivariogram.

At this point, we are ready to specify a kriging model. In this application, we employ an ordinary kriging (OK) model. In the ordinary kriging model, we assume $Z(\cdot)$ is intrinsically stationary, and the ordinary kriging predictor is a weighted average of the data points,

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i). \tag{5}$$

The weights, λ_i, are determined from the empirical or estimated semivariogram while imposing an unbiasedness and minimum mean-squared prediction error (MSPE) criterion. The unbiasedness criterion $(\mathrm{E}([\hat{Z}(s_0)]) = \mathrm{E}[Z(s_0)])^2$ implies that

$$\sum_{i=1}^{N} \lambda_i = 1. \tag{6}$$

To minimize the $\mathrm{MSPE} = \mathrm{E}[\hat{Z}(s_0) - Z(s_0)]^2$, we solve this constrained optimization problem via the method of Lagrange Multipliers with the unbiasedness constraint (6). The resulting optimal weights ensue, and these are used in (5) to arrive at the ordinary kriging predictor. The ordinary kriging predictor has the smallest mean-squared prediction error in the class of all linear unbiased predictors. That is, the ordinary kriging predictor is BLUP (best linear unbiased predictor).

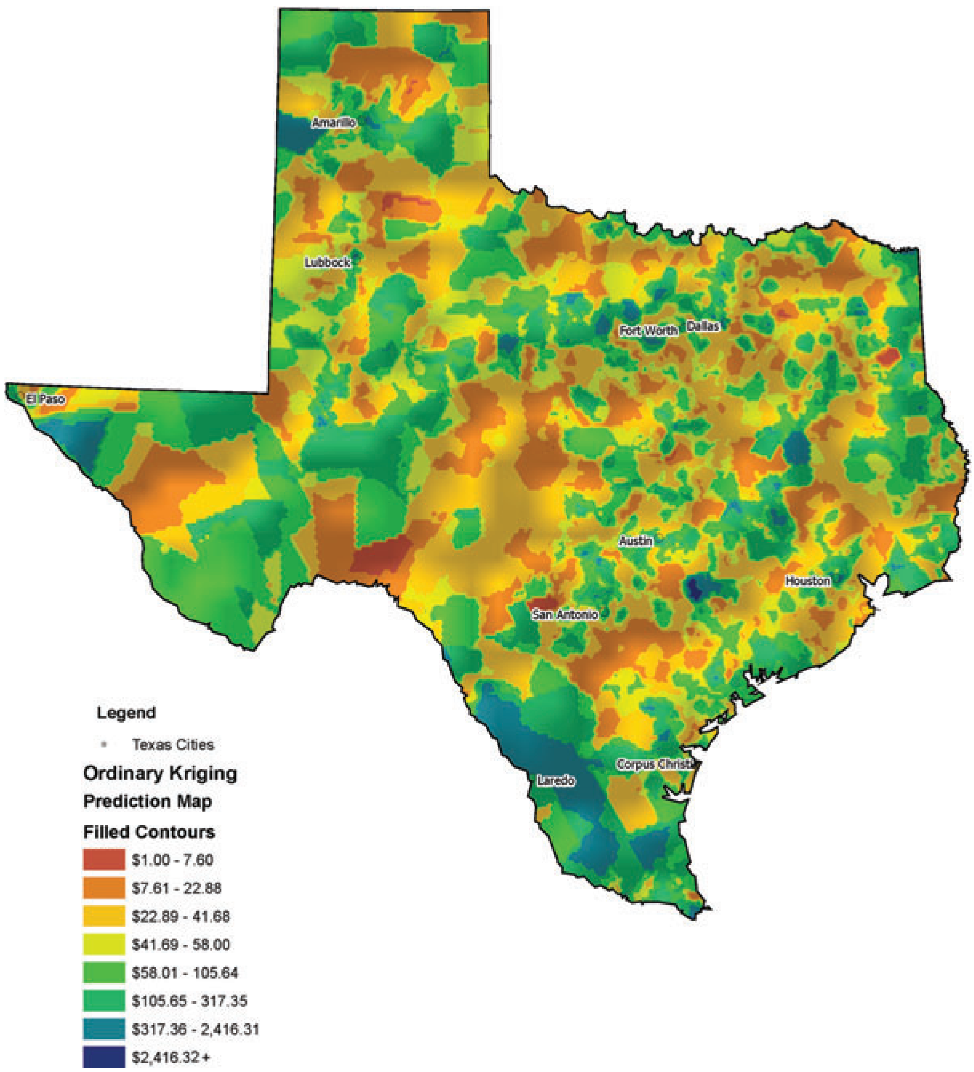
Finally, kriging is performed locally using search neighborhoods, and so we need to choose the number of neighbors from which to calculate kriging predictions. Here, we again rely on an empirical examination of the semivariogram as well as theoretical considerations. Our empirical semivariogram indicates that spatial dependency disappeared after about seven miles. Theory and previous research suggest that campaign fundraising networks are dense over a small number of neighborhoods, but quickly dissipate beyond a few miles (Cho 2003; Gimpel, Lee, and Kaminski 2006). Consistent with these findings, we employed a 20-neighbor criterion, with a minimum of four neighbors used to calculate weights for the most geographically isolated observations. Lastly, we used a log-transformation of contribution totals because the distribution of campaign contribution amounts exhibits positive skew.

Predictions and Cross-Validation

The kriging prediction map is shown in Figure 4. Major Texas cities are identified on the map to facilitate orientation and understanding. This map identifies the contributions one would expect to roll in from various areas in the state based on our kriging predictions. One can see how a campaign fundraiser might use this type of map to plot strategy: including identifying locations for events or developing prospect lists from zip codes that lie within fields predicted to contribute large sums. Perhaps

[7] Clearly, a semivariogram may be parameterized in many different ways, and the spherical family of functions is only one choice. It appears to be a good choice given the form of the empirical semivariogram and the fit shown by the fitted semivariogram model. In the modeling process, we also fitted parameters from other function families. Specifically, we used the K-Bessel family of functions to parameterize the semivariogram. This also provided a fairly good fit to the empirical semivariogram values, and the substantive results were very similar. Since the spherical model fit to the empirical semivariogram was a bit better, we chose to report the results that utilized the spherical family of functions.

FIGURE 4 Ordinary Kriging Prediction Map for Texas Fundraising



more importantly, the campaign can learn where *not* to prospect—the many locations that can be predicted to give very meager amounts. These areas are more efficiently cultivated via the Internet or direct mail as opposed to in-person events. In general, larger donations can be seen in Figure 4 to originate primarily in the state's metropolitan areas—centers of affluence—and particularly Dallas-Fort Worth, Houston, and Austin. Nevertheless, there are locations that are associated with neither metro areas nor affluence that could easily be missed relying upon "common sense" alone.[8]

The kriging results presented in Table 1 also provide useful information. For instance, the range parameter suggests that spatial autocorrelation dissipates beyond 7.4 miles. This may signal that there are limits on how far prospects are willing to travel to a fundraising event, consistent with the classic gravity model of spatial interaction.[9] Perhaps in most Texas locations, campaign finance staff can expect turnout from invitees within about a seven-mile radius of the event, but that invitations sent to households beyond this range will be accepted almost purely at random. We do note that this range is a global average for the entire geographic landscape. So this is a "general" rule, not necessarily meant to apply to a specific

[8]We note that the earlier anecdote about Laredo illustrates this point. Dallas should be an obvious hot spot for donations, but the potential may be insufficiently tapped if the donor networks are not identified and understood.

[9]Gravity models predict that behaviors related to spatial interaction such as migration and trade will mimic gravitational interaction as described by Newton's law of gravity. For instance, the gravity model of trade states that the volume of trade is an increasing function of the national income of trading partners and a decreasing function of the distance between the partners.

TABLE 1 Semivariogram Parameters and Cross-Validation Diagnostics for Model of Logged Campaign Contributions

Spherical γ	
Nugget	2.83
Sill	3.84
Range (in miles)	7.4
Partial Sill	1.01
Lag Size (in miles)	1
N lags	15
Cross-validation	
RMSPE	1.43
RMSE	1.04

instance or location. To identify the distance at which the spatial autocorrelation dissipates for a particular locale, one needs to examine the specific neighborhood in more detail. In identifying the optimal geographic allocation of event sites, this information may be of considerable value to event planners. The range will obviously differ from application to application and gives an indication of the spatial magnitude of "neighborhood effects." In some areas of the country, this value could encompass large areas while in others, similarities in behavior may be of more limited scope.

Since the range in the semivariogram model is an indicator of the extent of spatially related behavior, it also provides important information for those who might conduct survey research on contributors. Specifically, if we were conducting a survey of Texas donors concerned with the amount of money they contributed or were willing to contribute, we would not expect a randomly chosen donor to represent any other donors beyond a 7–8 mile surrounding area. The semivariogram therefore stands to be a useful guide to sampling strategy.

The standard errors associated with each prediction are mapped in Figure 5. The estimated variance depends on the number and proximity of the data points, the degree of spatial dependency in the data as quantified by the semivariogram, and provides a measure of interpolation uncertainty (Robertson 1987).

The standard errors are relatively small across the entire state, but especially in East and Central Texas. The bulk of the state's population, and the majority of campaign contributors, reside in this region bounded by Dallas in the north, Louisiana in the east, Houston and San Antonio to the south, and Interstate 35 on the west. For this vast and populous region, then, we have evidence that our predictions are fairly precise.

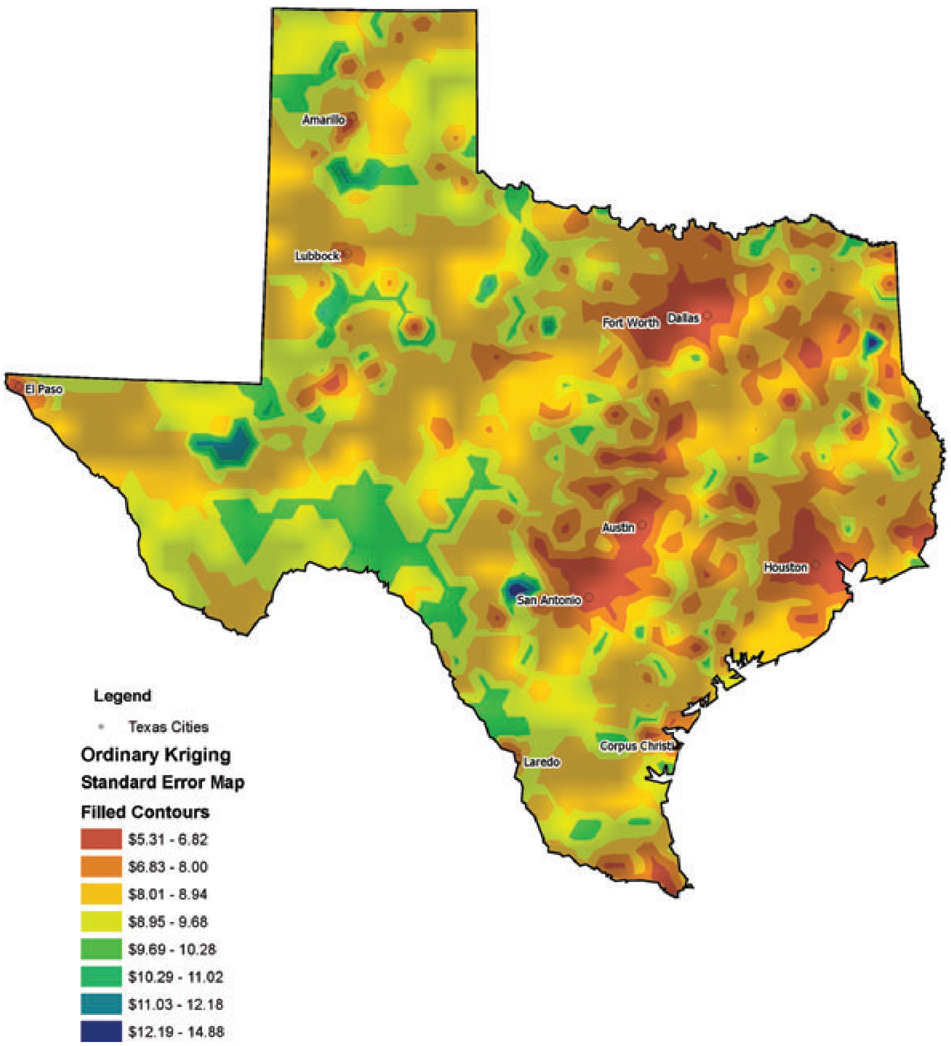
One way to assess the prediction model is through the technique of leave-one-out cross-validation (LOOCV). In LOOCV, we leave out one of the n observed points and use the remaining $n - 1$ points to predict its value. We repeat this process for each data point. The error at each data point is the difference between its observed and predicted values. The mean-squared prediction error is the mean of all of the squared differences between the observed values and those predicted from the other $n - 1$ points. A small mean-squared prediction error indicates that the predictions based on other points are close to the actual value, giving credibility to our prediction model. A measure of the overall error is the root mean-squared prediction error (RMSPE). The smaller the RMSPE, the better the predictive accuracy. The RMSPE for our data is 1.43, again giving us some confidence in our prediction model. We note that "precision" here is simply in terms of how well the semivariogram fit the empirical data, since the training set and the test set both originate in the same data set. We are not assessing predictions by obtaining new data and comparing those new data to the predicted contributions.

North Texas

It may be that an examination of such a large state presents an overwhelming amount of data and that the results are also too general to provide concrete guidance. If so, one could easily focus on a smaller metro area. Consider the "North Texas" area, a region that includes the large cities of Dallas, Fort Worth, and Arlington, and their surrounding suburbs, along with the outlying towns of Denton, Richardson, Euless, Garland, and Duncanville. *TRP* campaign records reveal that this region contributed 2,196 separate donations through March of the 2006 cycle (15.8% of total contributions at that date). Given the large sums that originate in this area, it is clearly an important location for prospecting. Results are presented in Table 2 and indicate that the prediction model for this area is not particularly unique in the state. If we had applied the general lessons from our examination of the entire state, we would not be far off. In addition, the results from cross-validation indicate that the smaller sample size and limited area have not undermined the precision of the estimates from the kriging model.

For a highly specific picture of contributor hot spots, we can turn to Figure 6, which shows the predictions for North Texas. The best areas for prospecting are shown in darker green and blue-green shades. While these contours do reflect areas of population density, they are more closely tied to affluence indicators. Dallas contributes more than Fort Worth, and specific low-density areas of Dallas, including Highland Park, can be counted on to

FIGURE 5 Ordinary Kriging Variance Map for Texas Fundraising



be superb locations for fundraising events and donation activity. The cities of Irving and low-density suburban development of Keller are good fundraising locations (see Figure 6). Duncanville, Garland, and Richardson, on the other hand, while suburbs of substantial size, cannot be counted on to contribute much. The map's contours indicate where contributions should be mined and what areas can be written off. Figure 7, the prediction error map, shows that the farther away from the observed locations, the larger the prediction errors become. That is, uncertainty (and hence the error measure) increases with increasing distance from the sampled point, until information from another sample point that is being approached begins to decrease the uncertainty again. Surely there are a few good prospects on the outskirts of the Dallas-Fort Worth metroplex, but the risk of reaping no return on a prospecting investment is far greater.

Discussion

Generally, kriging is a potentially useful technique for any politically relevant data that are regionalized (i.e., distributed in space). Our specific application has shown that because of the pronounced spatial concentration of campaign donors, prospect lists can be built and donor bases can be expanded by focusing attention on locations that have contributed in the past and exploring the surrounding neighborhoods. Certainly it is possible that a lonely prospector may strike a vein of ore off the beaten track, far away from existing deposits, but those strikes are likely to be few and far between and cannot be counted on to occur with much more than random frequency. Instead, a campaign prospector should not forge ahead randomly, but in specific locations clearly identified by a sampled set of previous contributors.

TABLE 2 Semivariogram Parameters and Cross-Validation Diagnostics for North Texas

Spherical γ	
Nugget	3.16
Sill	3.90
Range (in miles)	9.25
Partial Sill	0.74
Lag Size (in miles)	1
N lags	15
Cross-validation	
RMSPE	1.47
RMSE	0.997

To be sure, the wealth of knowledge possessed by an experienced fundraiser is acquired from years of field work, scheduling events, building relationships, examining lists of contributors, asking for donations on the telephone, and observing candidates beg for money (Smith 2005). This hard work can never be replaced by technical innovation, and there is no doubt that the experienced finance director or candidate will possess a strong intuitive sense of where to best prospect for new donations. At the same time, fundraising expertise is in limited supply and is usually quite costly. So, tools that can reduce these costs should be welcome in a world where the inflationary pressures associated with competitive campaigns relegate the substance of campaign politics to the sidelines. One way to make up considerable ground for those less experienced is through the methods we have introduced here. Without the enormous startup cost, one can use kriging methods to tap into the vast expanse of knowledge held captive by experienced fundraisers. For instance, the estimate of the range in kriging models helps to identify the extent of spatial autocorrelation or the confines beyond which geography ceases to function as a predictor of behavior. This distance may not coincide with traditional geographic boundaries that define cities or suburbs, and so they are not obvious. Kriging has the potential to provide a comprehensive picture of the geography of contributions

FIGURE 6 Ordinary Kriging Prediction Map for North Texas

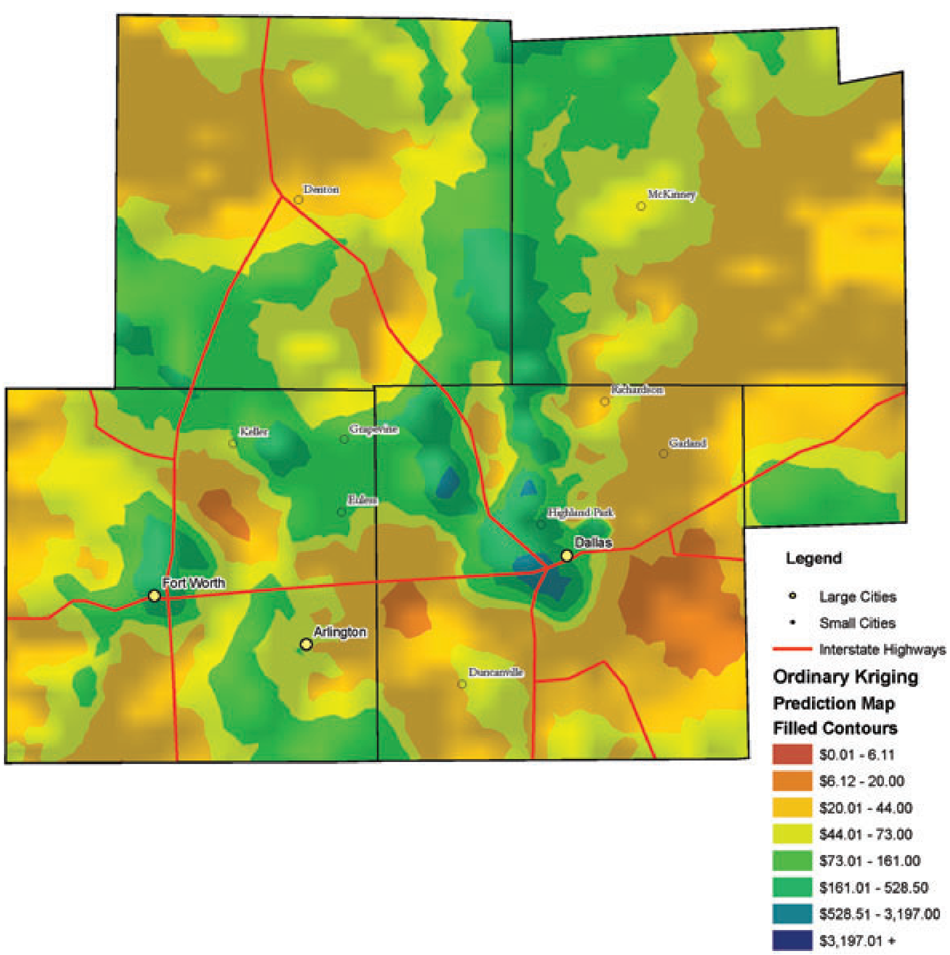
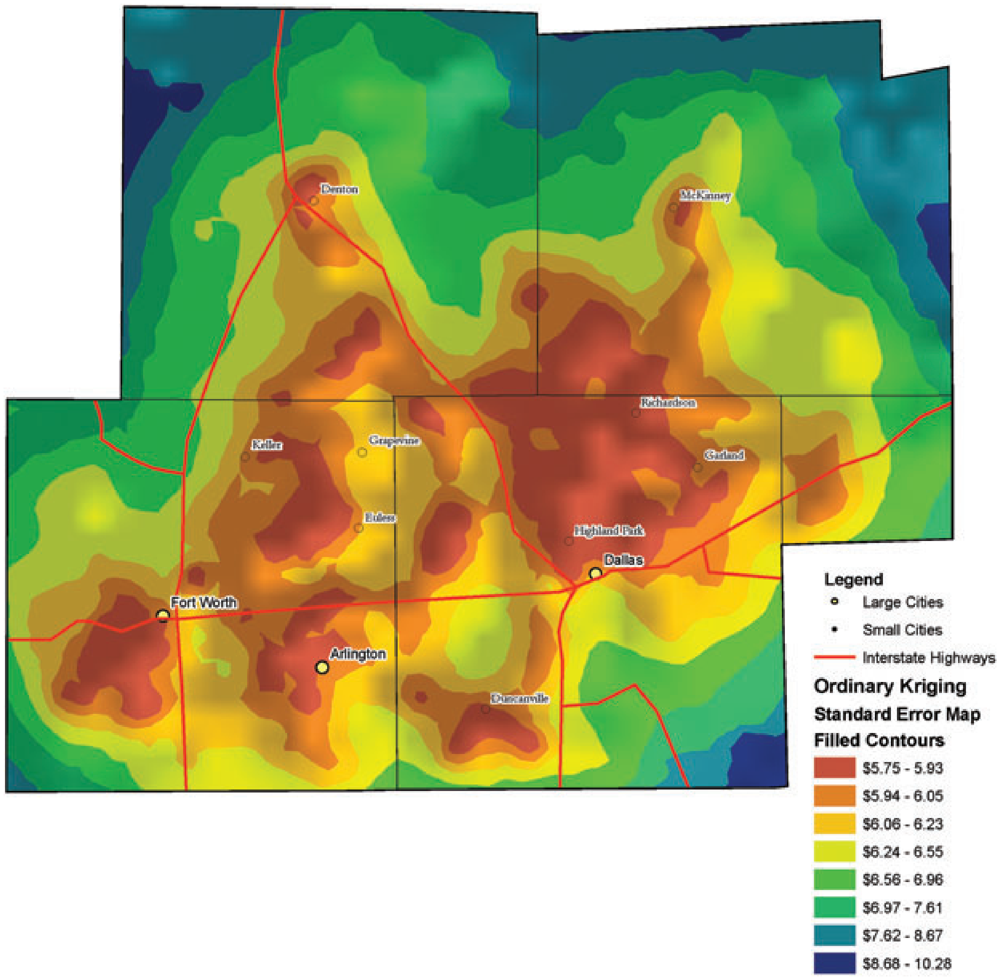


FIGURE 7 Ordinary Kriging Variance Map for North Texas



without investing the enormous capital that has traditionally been required.

Kriging is not the only method available for targeting donor prospects, and we do not present it in this light. For years now, contemporary marketers have capitalized on extensive data on neighborhood demographics and individual product preferences. For instance, the Republicans maintain the "Voter Vault" database, which journalists have touted as an effective weapon in several election successes for the GOP. Strategists and party activists use such microtargeting databases to model the propensity to vote and are now trying to model the propensity to donate with information about individuals' hobbies, magazine subscriptions, and other personal preferences.

Kriging provides several distinct advantages over such an approach. First, such databases are regarded as closely guarded trade secrets and are not available widely. So, we have again visited the land where the campaign playing field is hardly level. More pointedly, however, the fundamental concepts underlying contemporary marketing approaches and spatial models differ. Microtargeting begins with the assumption that potential contributors act atomistically. That is, their decision to contribute depends on their individual attributes (whether they drive a BMW, if they subscribe to *U.S. News and World Report*, their favorite brand of toothpaste, etc.). Kriging, on the other hand, eschews this tenet; contributors are *not* isolated actors. Rather if donors are influenced by other actors, and there appears to be little debate that they are, then there are clear benefits to utilizing techniques like kriging where the spatial component of this behavior is recognized and modeled. Moreover, this technique is not strictly limited to the experienced candidate who already has a donor base. Rarely do quality challengers begin with nothing. Even if they have never run for office before, they often have donor lists that they have borrowed or purchased from others. There are no guarantees of fundraising success using any extant method, but the clues offered here are plainly superior to guessing or a random walk through the phonebook.

Nonetheless, important limitations of this study and its method should be noted. For instance, we cannot conclude from these results that fundraising networks for all candidates would exhibit the same spatial pattern, nor can we assume that the spatial pattern would be the same for some other officeholder, a nonincumbent, or a candidate raising money in another state. Each case is unique though one may certainly notice strong patterns across similar campaigns, whether that similarity takes the form of party, level of office, race, or some other feature. Prediction is not explanation, and we do not propose to conflate them or to overlook the limitations of a prediction framework. In addition, spatial prediction, like spatial regression, and indeed like ordinary regression models, is unable to identify causal mechanisms. Instead, they simply identify imprints in the data that are consistent with the hypotheses proposed.

The conceptualization of space that we employ in this article has a critical influence on the understanding of spatial effects on contribution activity. Viewing space as a continuum is generally superior to viewing it as fragmented into predefined pieces unrecognizable to the person on the street. The kriging model captures not only the magnitude of spatial variation but also the scale of such variations (as reflected by the range of the semivariogram). Our results indicate that geographical coherence in contribution activity lies well below the scale of a large municipality, but probably above the scale of an average-sized zip code. With straightforward methods, one can gain a sense of the real extent of neighborhoods, community structure, and the fundraising potential at highly specific locations. Moreover, kriging may be carried out repeatedly, as new contributions arrive, with the objective of evaluating performance and detecting change (Rogerson and Sun 2001). Obtaining this grasp of the terrain is more than half the battle in an aspect of campaign politicking where neighborhoods and networks reign supreme.

We conclude by suggesting some avenues for future work. First, it is important to know if kriging could work to predict the geographic distribution of small ($<$\$250) as opposed to large contributions. Small contributors may be more responsive to direct mail appeals than large contributors. Our prima facie expectation is that the small contributor is far more dispersed and less network dependent than the large contributors. Additional investigations are needed before one can render a judgment on this score. Second, might we find that just as donor behavior is spatially autocorrelated, voluntarism is too. Money is a valuable input, but so is the time of loyal volunteers. Mapping the geographic distribution of current voluntary participation may lead to a greater understanding of the extent of recruitment networks.

Campaign contributing is an intriguing and important aspect of political participation, and one that has long interested political scientists. All contemporary signs point to the vast expansion of the contributor base for both parties in both state and national elections. Although we have yet to identify all of the complex mechanisms that motivate campaign giving, we can be confident that a vital component of any explanation is social and spatial in nature.

Appendix

Data

The data for this analysis is based on the complete donor records from the *Texans for Rick Perry* political campaign, obtained in April 2006, with permission of the campaign management and the candidate. These records included all donors who had contributed any amount to the *Texans for Rick Perry* campaign organization since January 2003. The records contained each amount contributed, the residential street address of the donor, along with the city, state, and five-digit zip code. We geocoded the data to add latitude/longitude coordinates that identify the residential street address of each donor on the surface of the earth. The latitude/longitude coordinates were pinpointed using a street network database produced by Geographic Data Technology Corporation, Lebanon, New Hampshire. The geographic information system (ArcGIS™, in this case) was used to place each donor on a map using these reference street ranges and the donor's residential street address.

The authors wish to thank Governor Rick Perry, David Carney, Luis Saenz, Leslie Sullivan, and Krystle Kirchmeyer for providing the campaign donation data used in this research, and sharing their insights about the campaign fundraising process.

References

Agnew, John A. 1990. From Political Methodology to Geographical Social Theory? A Critical Review of Electoral Geography. In *Developments in Electoral Geography*, ed. R. J. Johnston, F. M. Shelley, and P. J. Taylor. London: Routledge, pp. 152–221.

Ansolabehere, Stephen, John M. deFigueiredo, and James M. Snyder. 2003. "Why Is There So Little Money in U.S. Politics?" *Journal of Economic Perspectives* 17 (1): 105–30.

Brown, Clifford W. Jr., Linda W. Powell, and Clyde Wilcox. 1995. *Serious Money: Fundraising and Contributing in Presidential Nominating Campaigns*. New York: Cambridge University Press.

Cho, Wendy K. Tam. 2003. "Contagion Effects and Ethnic Contribution Networks." *American Journal of Political Science* 47 (2): 368–87.

Cressie, Noel A. C. 1989. "Geostatistics." *The American Statistician* 43 (4): 197–202.

Cressie, Noel A. C. 1993. *Statistics for Spatial Data*. New York: John Wiley and Sons.

Duncan, O. D., and B. Duncan. 1955. "A Methodological Analysis of Segregation Indexes." *American Sociological Review* 20: 210–17.

Francia, Peter L., John C. Green, Paul S. Herrnson, Lynda W. Powell, and Clyde Wilcox. 2003. *The Financiers of Congressional Elections: Investors, Ideologues and Intimates*. New York: Columbia University Press.

Freeman, Richard B. 1997. "Working for Nothing: The Supply of Volunteer Labor." *Journal of Labor Economics* 15 (1): 140–66.

Gerber, Alan, Donald P. Green, and Ron Shachar. 2003. "Voting May Be Habit Forming: Evidence from a Randomized Field Experiment." *American Journal of Political Science* 47 (3): 540–50.

Gimpel, James G., Frances E. Lee, and Joshua Kaminski. 2006. "The Geography of Campaign Contributions in American Politics." *Journal of Politics* 68 (3): 626–39.

Green, Donald P., and Ron Shachar. 2000. "Habit Formation and Political Behaviour: Evidence of Consuetude in Voter Turnout." *British Journal of Political Science* 30 (4): 561–73.

Herrnson, Paul S. 2006. *Congressional Elections: Campaigning at Home and in Washington*. Washington, DC: Congressional Quarterly Press.

Hollings, Ernest F. 2006. "Stop the Money Chase: A Constitutional Amendment Could Let Senators Be Senators." *Washington Post Opinion Editorial*. February 19 p. B07.

Huckfeldt, Robert, and John Sprague. 1992. "Political Parties and Electoral Mobilization: Political Structure, Social Structure and the Party Canvass." *American Political Science Review* 86 (1): 70–86.

Institute for Politics, Democracy and the Internet. 2006. *Small Donors and On-Line Giving: A Study of Donors to the 2004 Presidential Campaigns*. Campaign Finance Institute and George Washington University Graduate School of Political Management.

Jones, Ruth S., and Anne H. Hopkins. 1985. "State Campaign Fund Raising: Targets and Response." *Journal of Politics* 47 (2): 427–49.

Journel, A. G., and C. J. Huijbregts. 1978. *Mining Geostatistics*. London: Academic Press.

Ketchum, Carlton G. 1938. "Political Financing, the 1937 Model." *Public Opinion Quarterly* 2 (1): 135–40.

Kirchmeyer, Krystle. 2006. Personal Interview with Campaign Finance Director for Texans for Rick Perry. Transcribed notes from telephone interview.

Krasno, Jonathan S., Donald Philip Green, and Jonathan A. Cowden. 1994. "The Dynamics of Campaign Fundraising in House Elections." *Journal of Politics* 56 (2): 459–74.

Lazarsfeld, Paul F., and Robert K. Merton. 1954. Friendship as a Social Process: A Substantive and Methodological Analysis. In *Freedom and Control in Modern Society*, ed. M. Berger. New York: Van Nostrand Publishing Company, pp. 18–66.

Matheron, Georges. 1963. "Principles of Geostatistics." *Economic Geology* 58: 1246–66.

McBratney, A. B., and R. Webster. 1986. "Choosing Functions for Semi-Variograms of Soil Properties and Fitting Them to Sampling Estimates." *Journal of Soil Science* 37: 617–39.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: The MIT Press.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homopily in Social Networks." *Annual Review of Sociology* 27: 415–44.

Morrill, R. L. 1991. "On the Measure of Geographic Segregation." *Geography Research Forum* 11: 25–36.

Morrill, R. L. 1995. "Racial Segregation and Class in a Liberal Metropolis." *Geographical Analysis* 27: 2–41.

O'Loughlin, John, and Luc Anselin. 1996. "Geo-economic Competition and Trade Bloc Formation: United States, German, and Japanese Exports, 1968–1992." *Economic Geography* 72: 131–60.

O'Loughlin, John, Colin Flint, and Luc Anselin. 1994. "The Political Geography of the Nazi Vote: Context, Confession and Class in the 1930 Reichstag Election." *Annals Association of American Geographers* 84: 351–80.

Robertson, G. Phillip. 1987. "Geostatistics in Ecology: Interpolating with Known Variance." *Ecology* 68 (3):744–48.

Rogerson, Peter, and Y. Sun. 2001. "Spatial Monitoring of Geographic Patterns: An Application to Crime Analysis." *Computers, Environment and Urban Systems* 25 (6):539–66.

Shin, Michael, and John Agnew. 2002. "The Geography of Party Replacement in Italy, 1987–1996." *Political Geography* 21: 221–42.

Smith, Adam C. 2005. "GOP Embracing Young Lions of Fundraising." *St. Petersburg Times*. January 18 p. 1A.

Squire, Peverill, and John R. Wright. 1990. "Fundraising by Nonincumbent Candidates for the U.S. House of Representatives." *Legislative Studies Quarterly* 15 (1): 89–98.

Starr, Harvey, and Benjamin A. Most. 1983. "Contagion and Border Effects on Contemporary African Conflict." *Comparative Political Studies* 16: 92–117.

Steen, Jennifer A. 2006. *Money Isn't Everything: Self-Financing Candidates in Congressional Elections*. Ann Arbor: University of Michigan Press.

Stein, A., and L. C. A. Corsten. 1991. "Universal Kriging and Cokriging as Regression Procedures." *Biometrics* 47 (2): 575–87.

Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, MA: Harvard University Press.

Waller, Lance A., and Carol A. Gotway. 2004. *Applied Spatial Statistics for Public Health Data*. New York: John Wiley and Sons.

Ware, Alan. 1992. "Activist-Leader Relations and the Structure of Political Parties: 'Exchange' Models and Vote-Seeking Behaviour in Parties." *British Journal of Political Science* 22 (1): 71–92.

Webster, R. 1985. "Quantitative Spatial Analysis of Soil in the Field." *Advances in Soil Science* 3: 1–70.

Wong, David W. S. 1998. "Measuring Multiethnic Spatial Segregation." *Urban Geography* 19 (1): 77–87.

Wong, David W. S. 1999. "Geostatistics as Measures of Spatial Segregation." *Urban Geography* 20 (7): 635–47.

Wong, David W. S. 2004. "Comparing Traditional and Spatial Segregation Measures: A Spatial Scale Perspective." *Urban Geography* 25 (1): 66–82.