Ecological Inference

New Methodological Strategies

Edited by

Gary King
Harvard University

Ori Rosen
University of Pittsburgh

Martin A. Tanner
Northwestern University

5 Ecological Regression and Ecological Inference*

Bernard Grofman and Samuel Merrill

ABSTRACT

We propose three methods of ecological inference that guarantee feasible solutions but are simpler to implement than the method of King (1997). Each procedure provides estimates at the level of the ecological unit as well as a more aggregated level. The first method uses a simple squared distance minimization algorithm on the tomographic line segments. The second also generates a distance minimization, but in a space keyed to the slopes and intercepts of possible regression lines. The third determines the regression line that minimizes the sum of the areas between it and pairs of constraint line segments that are generated by a variant of the Duncan–Davis method of bounds. The procedures are implemented on an Excel spreadsheet and are available over the Internet. We present empirical applications, for which the first and third methods yield results that are quite similar to those produced by King's algorithm.

5.1 INTRODUCTION

In ecological inference we seek to make use of data that is aggregated at the level of ecological units to make inferences about the behavior of individuals. The ecological fallacy (Robinson, 1950) occurs when relationships between variables that obtain at the aggregate level are not found at the individual level. In this chapter three methods of ecological inference are proposed that are simpler than the sophisticated statistical models offered by King and his colleagues.

King estimates the unknown parameters of interest using maximum likelihood estimation (MLE) methods on a truncated bivariate normal or Beta distribution overlaid on the tomographic lines in (β^b, β^w) space. The first of our three methods uses a simple squared distance minimization algorithm on the tomographic lines. The district level solution is that point on the district tomographic segment that minimizes the (weighted) sum of the squared distances to the feasible tomographic line *segments* for the ecological units. For each unit the estimates for the parameters of interest are the coordinates of the nearest point on the unit tomographic line segment to the district solution. The analytic solution we obtain from this method is very similar to that of the Goodman regression model, but – unlike regression – our approach guarantees feasible solutions at both the precinct and the district level.

The second and third methods we propose can each be thought of as forms of *constrained* Goodman ecological regression. The first of these latter methods generates a

distance minimization in (m, b) space (where m and b are the slope and intercept of possible regression lines) rather than in (β^b, β^w) space. The last method operates in the original (X, T) space and finds the regression line that minimizes the sum of the areas between it and pairs of constraint line segments that are generated by a variant of the Duncan–Davis method of bounds.

These methods demonstrate that the contrast between ecological regression in the form proposed by Goodman (1953, 1959) and ecological inference of the sort described in King (1997) is too easily exaggerated. Each uses either King's extension of the Duncan–Davis (1953) method of bounds or a simple variant thereof. Each operates without any assumptions about the distribution of parameters, but bootstrap standard errors can be obtained to assess the results.

We compare the results of our methods with that of King for several artificial and real data sets. Our methods are implemented in Excel spreadsheets, which are available on the websites http://www.cbrss.harvard.edu/events/eic/book.htm and http://course.wilkes.edu/Merrill/. Two of the three methods produce results that are, in general, very close to those produced by King's algorithms.

5.1.1 Background

Since the critiques of scholars such as Robinson (1950), the use of ecological methods to attempt to specify individual level behavior from data that is available only at the level of ecological units has been both uncommon in the social sciences and highly suspect. It is now well known that ecological methods can sometimes yield quite misleading estimates, even of apparently simple statistics such as correlations. There have been a variety of attempts to resuscitate the use of ecological methods, such as the efforts of Goodman (1953, 1959) and Duncan and Davis (1953) to provide ecological estimates a solid statistical footing. In particular, these methods have been adopted for use in the analysis of racial bloc voting data in legal challenges to districting plans brought under the Voting Rights Act or the Fourteenth Amendment (Grofman, 2000).[1] But it is only following publication of Gary King's (1997) seminal work on ecological inference that the use of aggregate data on ecological units for purposes of directly inferring (mean levels of) *individual* behavior among individuals (entities) sorted into dichotomous or polychotomous categories has been undergoing a renaissance in political science research.

King (1997) argues that his approach to ecological inference is superior to Goodman's classic form of ecological regression for a number of reasons. Most notably, it makes use of all the information available about the data and the bounds on feasible parameter values, and guarantees that all estimates of unobservable individual parameters will be consistent with the feasible values for those parameters at the level of the ecological units used for analysis. While it is widely accepted that King (1997) represents a major advance on earlier methods such as Goodman's bivariate approach to ecological regression, King's approach to ecological inference has also been subject to strong attacks by scholars in both political science

---

[1] In voting rights challenges to districting plans (brought under Section 2 or Section 5 of the Voting Rights Act of 1965 as amended in 1982, or brought directly under the Fourteenth Amendment), analysis of voting by race was a legally mandated component of any litigation. Because survey data on local (or even state) elections is rarely available, analyses of the relationship between aggregate level voting patterns in the elections (usually measured at the precinct level) and the racial characteristics of these aggregate units has been used to make inferences about how members of each race are voting (Grofman and Migalski, 1988; Grofman, 1992). Despite the general disrepute of ecological methods over the past three decades, one area where ecological methods have been used by necessity is in the analysis of patterns of racial voting.

(Cho, 1998, 2001; Cho and Yoon, 2001; Anselin and Cho, 2002) and statistics (Freedman, Klein, Sachs, Smyth, and Everett, 1991; Freedman, Ostland, Roberts, and Klein, 1999). These authors argue that there are circumstances where his methods will be either inconclusive or wrong and that errors in inference may go undetected by his diagnostics. Others have argued that King overstates the distinctness of his approach from that of more traditional estimation techniques (McCue, 2001).

5.1.2 Proposed Methods

Our primary focus in this paper is not on critiques of ecological inference methods, but on offering three new methods of ecological inference that are easy to explain and very easy to calculate, e.g., using just an Excel spreadsheet. We suggest that each of these methods, which uses either King's extension of the Duncan–Davis method of bounds or a simple variant thereof, has many of the same nice properties as the methods proposed by King (1997) and King, Rosen, and Tanner (1999).[2] We will demonstrate that the contrast between Goodman-style ecological regression and ecological inference in the style of King is not so great as may appear.[3]

For simplicity of exposition, we will only look at bivariate analyses of the sort that can be done using the basic version of King's EZI computer program. We illustrate ecological methods as they might be applied to ascertain patterns of voting behavior in biracial contests involving at least one white candidate and one black candidate, using aggregate data (gathered, let us say, at the precinct level).[4] Here, we wish to understand what proportion of each group's votes go to a candidate identified with their own group.[5] Of course, our results have a much broader applicability than to the specific context of racial bloc voting.

We establish notation similar to that specified in Chapter 1:
For the ith *precinct*, let

X_i = proportion of the voters that are black,
T_i = proportion of the vote that goes to the black candidate,
β_i^b = proportion of black voters who vote for the black candidate,
β_i^w = proportion of white voters who vote for the black candidate.

For the entire *district*, let

X = proportion of the voters that are black,
T = proportion of the vote that goes to the black candidate,

---

[2] For more on this point see Grofman and Merrill (2002) and Silva de Mattos and Veiga (this volume, Chapter 15).

[3] Although this stark contrast is not King's own view (personal communication, 2001), we believe that, in large part because of the way the contrasts are emphasized in King (1997), most who have read this book have viewed King's method of ecological inference and Goodman's approach to ecological regression as almost completely opposite in nature.

[4] We will assume for convenience that we are only dealing with two groups of voters and that these two groups are mutually exclusive and exhaustive. As noted earlier, we refer to them as "black" and "nonblack," with "white" as a synonym for "nonblack." We also present our analyses for situations in which there is only one minority candidate, but extensions to situations with more than one minority candidate are straightforward.

[5] Analysis of racial bloc voting patterns is a context where it has been argued that the likely problems of ecological inference are minimized (Grofman, 1991, 1993, 2000). In that context, empirically, most methods yield very similar estimates, and methods such as standard ecological regression, when correctly applied and interpreted, have withstood legal challenges as well as challenges by expert witnesses skeptical of their accuracy.

B^b = proportion of black voters who vote for the black candidate,
B^w = proportion of white voters who vote for the black candidate.

The organization of this chapter is as follows.

We illustrate the direct link between the Goodman approach to ecological regression and the Duncan–Davis method of bounds, and we provide a theorem that allows us to derive a condition under which the results of the King (1997) approach and the answer obtained by ecological regression will be identical.

Next, we specify three "new" methods of ecological inference that use straightforward minimization algorithms that can be solved simply, e.g., using the Solver function in an Excel spreadsheet. For certain special cases, we can provide closed-form analytic solutions for these methods. Each of these methods draws, either directly or in transformed form, on King's (1997) seminal idea of using the Duncan–Davis (1953) method of bounds to construct line segments on which all feasible values of the unknown parameters must lie.

The first of these new methods operates in the same (β^b, β^w) space as that of King (1997). Rather than using MLE methods involving a truncated normal distribution or the Beta (see King, Rosen, and Tanner, 1999) or some other distribution, we solve a simple distance minimization problem to obtain the best-fitting joint prediction of the mean values of B^b and B^w. We then look at the projections from that point to the precinct-specific constraint boundaries (line segments) in (β^b, β^w) space to determine the best estimates of the individual β_i^b and β_i^w values.

The last two of our new methods can be viewed as variants of the Goodman ecological regression approach. They first produce a best estimate of the overall best-fitting bivariate regression, which yields feasible district values of B^b and B^w, and then use proximity to that line to generate precinct-specific estimates of the slopes m_i and intercepts b_i for the best-fitting overall regression lines for each precinct, from which the β_i^b and β_i^w values can be inferred. Empirically, we compare the results of our methods with the results of King's methods for artificial and real data sets.

5.2 ANALYZING THE LINK BETWEEN INDIVIDUAL BEHAVIOR AND BEHAVIOR RECORDED AT THE AGGREGATE LEVEL

We first illustrate the Duncan–Davis method and the simplest form of King's (1997) ecological inference model with an eleven-precinct set of hypothetical data (see Table 5.1) for which all methods will give essentially the same answer.

In each precinct, by the accounting identity,

$$T_i = \beta_i^b X_i + \beta_i^w (1 - X_i), \tag{5.1}$$

and similarly, for the district as a whole,

$$T = \beta^b X + \beta^w (1 - X). \tag{5.2}$$

While X and T are observable, the parameters of real interest, namely, the proportion of blacks who support the black candidate and the proportion of whites who support the black candidate, which we denote using β's, are unobservable. The problem is to get from the values we do know to those that we want to know about. By using the identity given

Table 5.1 Hypothetical illustrative eleven-precinct data set

Precinct	X_i Prop. black among voters	T_i Black cand. vote share	min β_i^b Min. black vote for black cand.	max β_i^b Max. black vote for black cand.	min β_i^w Min. white vote for black cand.	max β_i^w Max. white vote for black cand.
1	0.05	0.225	0.00	1.00	0.18	0.24
2	0.1	0.25	0.00	1.00	0.17	0.28
3	0.2	0.30	0.00	1.00	0.13	0.38
4	0.3	0.35	0.00	1.00	0.07	0.50
5	0.4	0.40	0.00	1.00	0.00	0.67
6	0.5	0.45	0.00	0.90	0.00	0.90
7	0.6	0.50	0.17	0.83	0.00	1.00
8	0.7	0.55	0.36	0.79	0.00	1.00
9	0.8	0.60	0.50	0.75	0.00	1.00
10	0.9	0.65	0.61	0.72	0.00	1.00
11	0.95	0.675	0.66	0.72	0.00	1.00
OVERALL MEAN	0.5	0.45	0.00	0.90	0.00	0.90
PRECINCT MEAN	0.5	0.45	0.18	0.89	0.04	0.75
GROUP POP	0.5	0.45	0.35	0.82	0.09	0.55

in Equation 5.1 above, which links T_i and X_i, with β_i^b and β_i^w, and combining it with our knowledge that vote proportions *must* lie between 0 and 1 (no ifs, ands, buts, or maybes), it is easy to show that any given pair of precinct values (X_i, T_i) gives rise to linear constraints on the feasible values of the β_i^b and β_i^w values for that precinct.

To see how this works, consider a simple example. Let us look at the data from Precinct 7 in Table 5.1. In that precinct, we have $X_i = 0.6$ and $T_i = 0.5$. Now, since 60% of the voters in the precinct are black and the black candidate got only 50% of the vote, *at most* five-sixths of the black voters (= 50%/60%) voted for the black candidate. On the other hand, even if all the white voters voted for the black candidate, since only 40% of the voters are white, *at least* one-sixth [= (50% − 40%)/60%] of the black voters must have supported the black candidate. Similarly, it is mathematically possible that every single white voter voted for the black candidate, and it is also possible that none of the white voters did so. By using data only from this precinct, the bounds we get on feasible patterns of black voting in the precinct do tell us that (given the actual X_i and T_i values) we must have between one-sixth and five-sixths of the black voters in that precinct supporting the black candidate, but the proportion of white voters who supported the black candidate could be anywhere between 0% and 100%.

The Duncan–Davis (1953) method can be used to get tight bounds either in precincts that are racially homogeneous or in precincts that vote lopsidedly for a candidate of one race. We have shown in Table 5.1 the maximum and minimum values of β_i^b and β_i^w for each of the eleven precincts. It is apparent that, for Precinct 1, the most heavily white precinct, although the bounds on the black vote are not at all informative, we can pin down the proportion of white voters supporting the black candidate as falling between 18% and 24%. Similarly, for Precinct 11, the most heavily black precinct, although the bounds on the white vote are not at all informative, we can pin down the proportion of black voters supporting the black candidate as falling between 66% and 72%.
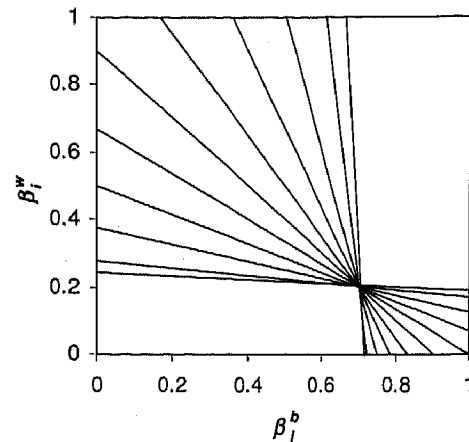
Figure 5.1. Tomographic plot for data in Table 5.1.

5.2.1 Tomographic Plots

When $X_i = 0.6$ and $T_i = 0.5$, not only is it true that $(\frac{1}{6}, 1)$ and $(\frac{5}{6}, 0)$ are feasible outcomes, but it is easy to see that all points on the line segment between the points $(\frac{1}{6}, 1)$ and $(\frac{5}{6}, 0)$ are also feasible, and are given by the equation

$$\beta_i^w = -\frac{3}{2}\beta_i^b + \frac{5}{4}$$

The only portion of this line that is of interest is the line segment containing the feasible values, i.e., the values on this line that lie at or between the points $(\frac{1}{6}, 1)$ and $(\frac{5}{6}, 0)$. We show in Figure 5.1 the precinct-based constraints on the joint (β_i^b, β_i^w) pairs for the data in Table 5.1. This type of joint constraint diagram, known as a *tomographic plot*, with values plotted in (β^b, β^w) *space*, was first introduced into the social sciences by King (1997) on p. 81 of his book, and used repeatedly thereafter.

Understanding what such a diagram shows is absolutely critical to understanding King's approach to the problem of reliable ecological inference and our own similar approaches. It is also critical to understanding alternative approaches such as the neighborhood model of Freedman et al. (1991).

Note that all the feasibility constraint lines in Figure 5.1 intersect at a single point (.70, .20). The parameter values at this intersection point correspond to the estimate for the mean value of the unobserved parameters we get from Goodman's ecological regression for the same data set. King's EZI algorithm also yields these same mean values for this data set, as it usually does when the tomographic segments meet at a point.

We now offer a simple result linking the values in the tomographic plots generated as the basis for King's method of ecological inference and the results of the classic Goodman ecological regression for the special case of the feasibility constraint lines intersecting at a single point.

Theorem 1. If the tomographic line segments used as the basis for King's ecological inference have a unique intersection, this intersection will be at a point, $(\hat{B}^b, \hat{B}^w)$, which corresponds to the (B^b, B^w) values derived from Goodman's method of bivariate ecological regression.

Proof. See Appendix 1.

Unfortunately, even when there is a unique intersection point of the tomographic plot lines, that intersection need not be within the unit square, i.e., need not be a feasible value.[6] Indeed, we might anticipate that, even in the absence of a unique intersection of the line segment bounds in the tomographic plot, when Goodman's ecological regression method yields a feasible estimate of mean (B^b, B^w) values, it is likely that the results of Goodman's approach and that of King's approach to ecological inference will not be far apart. The differences between the two approaches appear likely to arise when Goodman's ecological regression yields out-of-bounds estimates for one or more of the mean (or precinct specific) parameters. We will return to this issue, i.e., the circumstances under which different methods are likely to give rise to different answers, later in the chapter.

5.3 A SIMPLE DISTANCE MINIMIZATION ALGORITHM FOR ECOLOGICAL INFERENCE: METHOD I

It follows from equations (1) and (2) that

$$\beta_i^w = \frac{-X_i}{1 - X_i}\beta_i^b + \frac{T_i}{1 - X_i} \tag{5.3}$$

and

$$B^w = \frac{-X}{1 - X}B^b + \frac{T}{1 - X}. \tag{5.4}$$

In the example specified in Table 5.1, $B^w = -B^b + 0.9$, and similar equations hold for each precinct. Any pair of (B^b, B^w) values that lie on the district line segment specified in Equation 5.4 is compatible with the overall pattern of racial bloc voting in this data. But which point on this line segment is the most plausible estimate of this pair of values?

A simple approach is to look to see how far the various points on this line are from the other line segments in the tomographic plot. If, for example, there is a unique intersection of all the other line segments with each other, then the aggregate line-segment bound must also pass through that intersection. In this special case, it would seem that a very compelling case can be made for choosing the intersection point as our "best" estimate of the (B^b, B^w) values, at least if that point consists of jointly feasible values. In general, we can find the squared distance from each point on the aggregate line segment to each of the precinct line segments, and find the point that minimizes the (weighted) sum of those distances. Such distances will be interpreted later.

Our plan is to compute numerically – for each point on the overall tomographic constraint line defined by Equation (4) – the sum of the squared distances from that point to each of the precinct-level tomographic line *segments*. If the perpendiculars to the precinct-level tomographic lines intersect these lines at points in the feasible region, then a closed form solution for (B^b, B^w) can be derived (see Equations 5.13a and 5.13b below). If, instead, a perpendicular to a precinct-level tomographic line falls outside the feasible region, the shortest distance from a given point on the district line to the precinct-level segment must

---

[6] If, for example, the (X, T) values are (.3, .2), (.5, .5), and (.7, .8), the tomographic lines intersect at $\beta^b = 1.25$ and $\beta^w = -0.25$, which represent a point outside the unit square.

be computed to the nearer endpoint of the segment. The specifications of these endpoints P_1 and P_2 follow simple rules:[7]

$$\text{If } T_i \leq 1 - X_i \text{ then } P_1 = \left(0, \frac{T_i}{1 - X_i} \right);$$

$$\text{otherwise } P_1 = \left(\frac{T_i - (1 - X_i)}{X_i}, 1 \right). \tag{5.6a}$$

$$\text{If } T_i \geq X_i \text{ then } P_2 = \left(1, \frac{T_i - X_i}{1 - X_i} \right);$$

$$\text{otherwise } P_2 = \left(\frac{T_i}{X_i}, 0 \right). \tag{5.6b}$$

To implement this plan, it remains only to determine formulas for the points of intersection (to be used when they lie in the feasible region). As noted above, we have

$$\beta_i^w = \frac{-X_i}{1 - X_i} \beta_i^b + \frac{T_i}{1 - X_i} \tag{5.7}$$

as the equation for each precinct constraint line. If (B^b, B^w) lies on the aggregate constraint line, the line through this point and perpendicular to a precinct constraint line given by Equation 5.7 is given by

$$\beta_i^w = \frac{1 - X_i}{X_i} \beta_i^b + B^w - \frac{1 - X_i}{X_i} B^b. \tag{5.8}$$

The point of intersection of the precinct constraint line and this perpendicular is given by

$$\beta_i^b = \frac{X_i T_i - B^w X_i (1 - X_i) + B^b (1 - X_i)^2}{X_i^2 + (1 - X_i)^2} \tag{5.9}$$

and β_i^w can then be obtained from Equation 5.8.

In general, what we want to do is find the point on the district-level tomographic line that minimizes the sum of the squared distances from that point to all the line segments that define the precinct-specific joint bounds on the β_i^b and β_i^w values. First note that, from Equation 5.8,

$$\beta_i^w - B^w = \frac{1 - X_i}{X_i} \left(\beta_i^b - B^b \right),$$

---

[7] Note that the conditions on T_i in Equations 5.6a and 5.6b need not be complementary; it is the two conditions within 5.6a and within 5.6b that are complementary. In the degenerate case for which $X_i = 1$, if $T_i \leq 1 - X_i$ then $P_1 = (0, 1)$; if $T_i \geq X_i$ then $P_2 = (1, 0)$.

so that the square of the distance from (B^b, B^w) to the precinct constraint line, i.e., to the point of intersection given by Equation 5.9, is

$$
\begin{aligned}
d_i^2 &= \left(\beta_i^b - B^b\right)^2 + \left(\beta_i^w - B^w\right)^2 \\
&= \left(\beta_i^b - B^b\right)^2 \frac{X_i^2 + (1 - X_i)^2}{X_i^2}.
\end{aligned}
\tag{5.10}
$$

However, using Equation 5.9, we obtain

$$
\beta_i^b - B^b = \frac{X_i T_i - B^w X_i (1 - X_i) - X_i^2 B^b}{X_i^2 + (1 - X_i)^2}.
$$

Together with Equation 5.10, this implies that

$$
\begin{aligned}
d_i^2 &= \frac{\left[T_i - X_i B^b - (1 - X_i) B^w\right]^2}{X_i^2 + (1 - X_i)^2} \\
&= w_i^2 \left[T_i - X_i B^b - (1 - X_i) B^w\right]^2,
\end{aligned}
\tag{5.11}
$$

where the weights w_i are given by

$$
w_i = \frac{1}{\sqrt{X_i^2 + (1 - X_i)^2}}.
$$

Note that the distance d_i can be interpreted as the weighted difference between the proportion of voters for the black candidate in the ith precinct and what that proportion would be if the proportions voting for the black candidate broken down by race were given by B^b and B^w, that is, the same as in the district as a whole. Hence, it makes sense to seek values of B^b and B^w that would minimize the squares of these differences. In fact, the numerator in Equation 5.11 is $(T_i - \hat{T}_i)^2$, where $\hat{T}_i$ is the ith fitted value under Goodman regression.

If all points of intersection are in the feasible region, we simply minimize $\sum_i d_i^2$ subject to the constraint that B^b and B^w are feasible (lie on the district constraint line), i.e., that

$$
XB^b + (1 - X)B^w = T.
\tag{5.12}
$$

Solving this constrained optimization problem by Lagrange multipliers, we obtain two linear equations in B^b and B^w:

$$
B^b \sum_i w_i^2 X_i (X_i - X) + B^w \sum_i w_i^2 (1 - X_i)(X_i - X) = \sum_i w_i^2 T_i (X_i - X),
$$

$$
B^b X + B^w (1 - X) = T,
$$

which yield the solutions

$$B^b = \frac{\sum_i w_i^2 (X_i - X) [(1 - X) T_i - (1 - X_i) T]}{\sum_i w_i^2 (X_i - X)^2}, \qquad (5.13a)$$

$$B^w = \frac{\sum_i w_i^2 (X_i - X) [X_i T - X T_i]}{\sum_i w_i^2 (X_i - X)^2}. \qquad (5.13b)$$

Thus, in the special case in which all intersection points are in the feasible region, we have obtained closed-form solutions for B^b and B^w. These solutions are simple to compute on a spreadsheet and closely resemble the form of solutions to an ordinary least squares regression problem.[8] However, in solving our optimization problem, we are only interested in points of intersection (β_i^b, β_i^w) that specify *feasible* values for the respective precincts. Accordingly, if the point of intersection is outside the feasible region, we modify d_i^2 to be the squared distance to the nearer endpoint of the precinct line segment where it intersects the boundary of the feasible region. We then choose those values of B^b and B^w that lie on the district tomographic line and that minimize $\sum_i d_i^2$.

Standard errors and confidence intervals can be computed by a bootstrap method. This is done by repeated sampling with replacement from the data set, recomputing the parameter estimates, and determining the standard deviation of these estimates (see Efron and Tibshirani, 1993).

Each precinct-level estimate is the pair (β_i^b, β_i^w) that minimizes the expression $(\beta_i^b - B^b)^2 + (\beta_i^w - B^w)^2$. It is the intersection point of the perpendicular to the precinct tomographic line if this value is feasible, and otherwise is the nearest endpoint of the precinct tomographic line segment to the district solution point (B^b, B^w). These computations can be implemented in an Excel spreadsheet and are available on the websites `http://www.cbrss.harvard.edu/events/eic/book.htm` and `http://course.wilkes.edu/Merrill/` through Internet Explorer.

District parameter estimates for Method I are presented later for several artificial and real data sets in Tables 2–4; precinct-level estimates are given for one real data set in Table 3. These results are discussed in Section 5.5.

If not all precincts are of equal size, we weight the d_i^2 by the number N_i of voters in precinct i, i.e., we minimize $\sum N_i d_i^2$. Equations 5.13a and 5.13b are replaced by

$$B^b = \frac{\sum_i w_i^2 N_i (X_i - X) [(1 - X) T_i - (1 - X_i) T]}{\sum_i w_i^2 N_i (X_i - X)^2}, \qquad (5.14a)$$

$$B^w = \frac{\sum_i w_i^2 N_i (X_i - X) [X_i T - X T_i]}{\sum_i w_i^2 N_i (X_i - X)^2}. \qquad (5.14b)$$

---

[8] In this special case, the solution would be identical to the ordinary least squares solution if the weights w_i in Equation 5.13 were all identical.

5.4 EXTENDING THE DUNCAN–DAVIS METHOD OF BOUNDS TO DEVELOP TWO NEW FORMS OF GOODMAN'S ECOLOGICAL REGRESSION APPROACH: METHODS II AND III

King's ecological inference approach makes use of tomographic plots that constrain the values of unobservable individual-level parameters (β_i^b and β_i^w) to lie within feasible bounds for each of the ecological units. Ecological inference uses maximum likelihood methods to derive overall estimates of these unobservable parameters. We show that Goodman's approach to ecological regression can be adapted to make use of distance minimization methods that constrain the values of slope and intercept parameters so that the estimates of the unobservable individual-level parameters (β_i^b and β_i^w), and the mean values for those parameters, remain within feasible bounds. Indeed, we provide two different methods for doing so.

5.4.1 Adapting our Previous Distance Minimization Algorithm for Use in (m, b) Space Rather than (β^b, β^w) Space: Method II

Our first proposed integration of ecological inference and ecological regression uses a mathematical device to shift from the usual (X, T) space to a new space defined in terms of m and b, the slope and intercept parameters of the bivariate ecological regression. We derive the defining values for the line segments in that space from the Duncan–Davis (1953) method of bounds. Because

$$m = \beta^b - \beta^w$$

and

$$b = \beta^w,$$

we have

$$m = \left(b - \frac{T}{1-X} \right) \left(\frac{1-X}{-X} \right) - b = \frac{-b}{X} + \frac{T}{X}.$$

This expression may be rewritten as

$$T = mX + b.$$

For our example for which (X, T) = (.5, .45), the expression for the feasible overall line in (m, b) space is

$$m = -2b + 0.9.$$

Similar equations hold for each precinct.

Once we have constructed this set of equations, we apply the same methods as in the Section 5.3 to find the point on the line in the equation above that minimizes the (constrained) sum of squares to the various precinct-specific line segments. Only, now we are operating in (m, b) space rather than in (β^b, β^w) space.

The feasible region in (m, b) space is a diamond with corners at (0, 0), (0, 1), (−1, 1), and (0, 1). The squared distance from a point (m_0, b_0) on the district feasible line segment
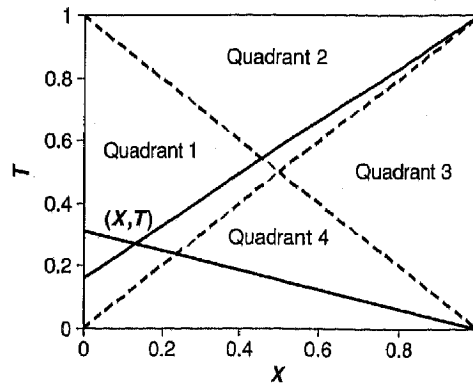
Figure 5.2. Quadrants for defining bounding lines for cones (dashed lines), with example of a pair of bounding lines through the point (X, T) (solid lines).

to a precinct feasible line segment is given by

$$d_i^2 = \frac{X_i^4}{(1 + X_i^2)^3}(b_0 + T_i - m_0 X_i)^2$$

if the foot of the perpendicular to the precinct line lies in the diamond, and otherwise by the distance to the nearest endpoint on that line segment. Because the transformation from the feasible square of (β^b, β^w) space to the diamond of (m, b) space alters distances, the minimization problems in Methods I and II are not identical. Simple examples show that precincts with symmetric patterns are treated symmetrically in (β^b, β^w) space, but not in (m, b) space.[9] As expected, Method II yields estimates for (B^b, B^w) that may be quite different from those obtained by King (1997) or by our Method I (or by our Method III below).

5.4.2 Operating Directly in (X, T) Space on the Set of Feasible Regression lines: Method III

Here we seek to build into the Goodman ecological regression approach the constraints on feasible values generated by the Duncan–Davis (1953) method of bounds. We now specify a pair of regression lines in the original (X, T) space that give bounds, for each ecological unit, for the jointly feasible values of m_i and b_i derived from the set of constraints on jointly feasible β_i^b and β_i^w values.

5.4.3 Defining the Cone of Feasible Values

Consider a point (X_i, T_i) that is an observed value for a given ecological unit (such as a precinct). If an ecological regression line passing through that point is to yield values of m and b that are feasible, it must be the case that the regression line intersects the line $X = 0$ somewhere between $T = 0$ and $T = 1$ and that it also intersects the line $X = 1$ somewhere

---

[9] For example, consider a district with three precincts with $(X_i, T_i) = (.3, .1)$, $(.5, .5)$, and $(.7, .9)$. The first and third precincts are symmetric in all respects and are equidistant from (and symmetric to) the district tomographic line segment in (β^b, β^w) space. In (m, b) space, however, the optimizing point on the district line is closest to an interior point on the feasible segment for precinct 1 but to an endpoint for the feasible segment for precinct 3. The distances involved are not the same.
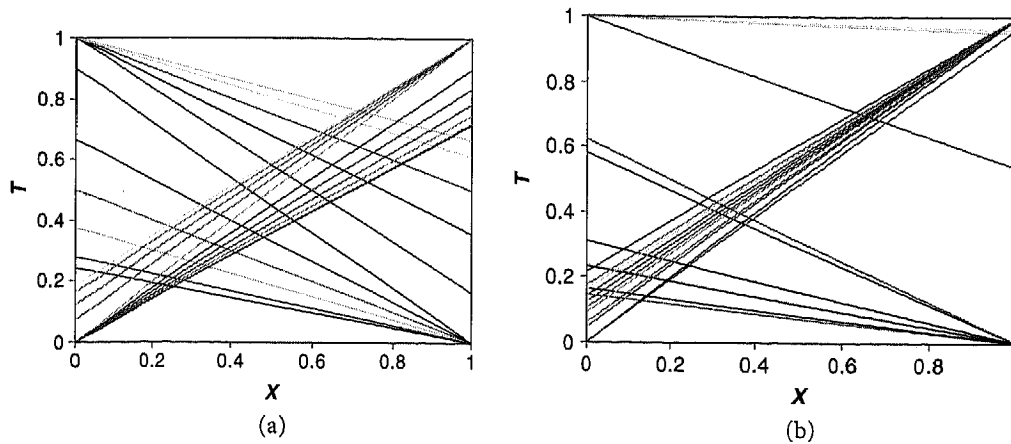
Figure 5.3. Bounding lines for cones of feasible values for 11 precincts for data from (a) Table 1, (b) Carter.

between $T = 0$ and $T = 1$. To see how these constraints apply, divide the unit square in standard (X, T) space into quadrants, as shown in Figure 5.2.

In quadrant 1, define a cone of feasible values passing through a point (X_i, T_i) by requiring that one defining (extremal) line of the cone pass through the point $(1, 0)$ and the other defining line pass through $(1, 1)$. Similarly, for points in quadrant 2, the defining lines of the cone must pass through $(0, 1)$ and $(1, 1)$; for points in quadrant 3 they must pass through $(0, 0)$ and $(0, 1)$; for points in quadrant 4 they must pass through $(0, 0)$ and $(1, 0)$. For points in quadrant 1, one of the two defining lines of the cone [that which passes through the point $(1, 0)$] must be a line whose m value is equal to $-T_i/(1 - X_i)$ and whose b value is equal to $T_i/(1 - X_i)$, while the other defining line [that which passes through the point $(1, 1)$] must be a line whose m value is equal to $(1 - T_i)/(1 - X_i)$ and whose b value is equal to $(T_i - X_i)/(1 - X_i)$.

In like manner, the cone of feasible regression lines for points in quadrant 2 is characterized by a bounding line with m value equal to $(1 - T_i)/(-X_i)$ and b value equal to 1, and a second bounding line with m value equal to $(1 - T_i)/(1 - X_i)$ and b value equal to $(T_i - X_i)/(1 - X_i)$. The cone for quadrant 3 is characterized by a bounding line with m value equal to T_i/X_i and b value equal to 0, and a second bounding line with m value equal to $(1 - T_i)/(-X_i)$ and b value equal to 1. Finally, the cone for quadrant 4 is characterized by a bounding line with m value equal to T_i/X_i and b value equal to 0, and a second bounding line with m value equal to $-T_i/(1 - X_i)$ and b value equal to $(T_i)/(1 - X_i)$. We can illustrate these boundary lines for the set of data in Table 5.1 (Figure 5.3a), and for an 11-precinct data set drawn from a real-world biracial legislative contest in the Deep South in the 1990s (Figure 5.3b; see also Section 5.5).

Our plan is to choose values of m and b that satisfy the district-wide constraint

$$T = mX + b,$$

i.e.,

$$m = (T - b)/X,$$

and minimize the sum of the areas between this line segment and the bounding line segments of the cones associated with the precinct values X_i and T_i. These areas are restricted to feasible values of X_i, i.e., $X_i \in [0, 1]$. The detailed calculations are given in Appendix 2.

An Excel spreadsheet to implement these calculations for Method III is available on the websites `http://www.cbrss.harvard.edu/events/eic/book.htm` and `http://course.wilkes.edu/merrill/`. Standard errors and confidence intervals can be constructed by bootstrap methods.

We can plot the (m, b) values for each pair of lines that constitute the extremal lines of the boundary cones as points in (m, b) space. For example, if (X_i, T_i) is in quadrant 1, then the two points are $(-T_i/(1 - X_i), T_i/(1 - X_i))$ and $((1 - T_i)/(1 - X_i), (T_i - X_i)/(1 - X_i))$. The line connecting these two points has equation $b = -X_i m + T_i$. In fact, the corresponding line for each of the other cones has exactly the same equation.

5.5 COMPARING THE KING ECOLOGICAL INFERENCE ESTIMATES AND THOSE OF OUR SIMPLIFIED APPROACHES

Aggregate parameter estimates produced by the basic version of King's (1997) MLE program and the estimates generated by our distance- or area-minimizing algorithms are typically similar to one another, but need not be identical. We show this in Table 5.2 for three hypothetical data sets (data sets A, B, and C).

The circumstances in which King's basic method and our least squares and area-minimizing approaches can be expected to give more divergent results occur when heavily truncated tomographic line segments pull the solution away from a more plausible

Table 5.2 Parameter estimates for B^b and B^w by alternative methods for artificial data sets

Data set	Parameter	King's method	Method I	Method II	Method III	Regression
A	B^b	.838 (.005)	.830 (.023)	.839	.828 (.022)	.833 (.020)
	B^w	.282 (.005)	.290 (.017)	.281	.292 (.015)	.287 (.020)
B	B^b	.713 (.053)	.748 (.066)	.580	.656 (.069)	1.500 (.000)
	B^w	.287 (.053)	.252 (.077)	.420	.344 (.086)	−.500 (.000)
C	B^b	.917 (.024)	.829 (.083)	.672	.937 (.135)	1.007 (.035)
	B^w	.083 (.024)	.171 (.098)	.328	.063 (.166)	−.007 (.035)

Data set A:

X	0.10	0.18	0.26	0.34	0.42	0.50	0.58	0.66	0.74	0.82	0.90
T	0.34	0.37	0.44	0.50	0.53	0.52	0.64	0.59	0.69	0.75	0.79

Data set B:

X	0.30	0.34	0.38	0.42	0.46	0.50	0.54	0.58	0.62	0.66	0.70
T	0.10	0.18	0.26	0.34	0.42	0.50	0.58	0.66	0.74	0.82	0.90

Data set C:

X	0.10	0.15	0.20	0.80	0.85	0.90
T	0.05	0.15	0.25	0.75	0.85	0.95

Note: Bootstrap standard errors are given in parentheses.

Figure 5.4. Tomographic plot for data set C. (Bold line is the district tomographic line.)
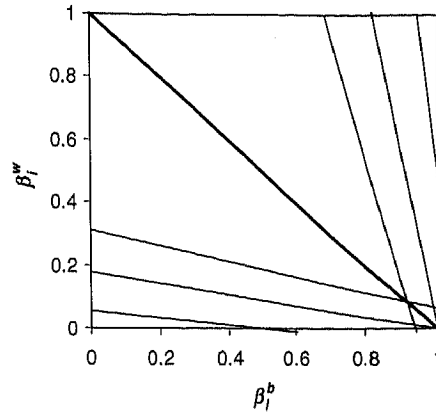
convergence of other, less-truncated tomographic line segments. A case in point is data set C (see Table 5.2 and Figure 5.4). Here the endpoints (.5, 0) and (1, .5) on tomographic line segments 1 and 6 have what appear to be inordinate effects on the parameter estimates for Method I. For this data set, Method III and King's method provide a more polarized but more plausible solution than Method I. The facts that not more than 50% of the blacks vote for the black candidates in Precinct 1 (in our example scenario) and not less than 50% of the whites vote for the black candidate in Precinct 6 may seem inconsistent with the rest of the district. Yet, given the small number of blacks in Precinct 1 and small number of whites in Precinct 6, such statistics may commonly occur due to random variation.

5.5.1 Weighting by Informativeness

Precincts that are mostly black are most informative in estimating B^b, whereas those that are mostly white are most informative in estimating B^w. Accordingly, we define a version of Method I *weighted by informativeness* by replacing the raw distance between (B^b, B^w) and a precinct tomographic line segment with a metric in which the coordinates are weighted by the proportions of blacks and whites. Thus, we define weighted squared distance

$$N_i \left[X_i \left(\beta_i^b - B^b \right)^2 + (1 - X_i) \left(\beta_i^w - B^w \right)^2 \right],$$

where N_i is the number of voters in precinct i, and where we recall that X_i and $1 - X_i$ are the proportions of blacks and whites, respectively, in the electorate. Given this weighting, if all minimizing points (β_i^b, β_i^w) do lie in the feasible region, the solution for (B^b, B^w) is identical with that of ecological regression.

For data set C, weighting the coordinates by the proportions of blacks and whites makes a substantial difference. The weighted estimates (see Table 5.4) for (B^b, B^w) are (.957, .043), in contrast to unweighted estimates of (.829, .171). King's estimates (.917, .083) are intermediate, and closer to our weighted estimates.

In practice, when we are looking at data from U.S. elections involving candidates of more than one race (especially data from the South), most methods are apt to yield values of at least the aggregate parameters B^b and B^w that are reasonably close to one another.[10] To

---

[10] In the specific context of racial bloc voting analyses, they are also likely to be not very different from the estimates generated by ecological regression, at least when those estimates are within (0,1) bounds (Grofman, 2000).

Table 5.3 Comparison of parameter estimates from methods I–III and King's method for the Carter 11-precinct data set

(a) District-level estimates for B^b and B^w

Data set	Parameter	King's method	Method I	Method II	Method III	Regression
Carter 11-precinct sample	B^b	.978 (.005)	.974 (.043)	.940	.964 (.059)	.974 (.026)
	B^b	.136 (.005)	.138 (.016)	.155	.144 (.016)	.138 (.017)

Carter data:

X	0.13	0.67	0.16	0.04	0.33	0.02	0.04	0.02	0.31	0.99	0.95
T	0.27	0.69	0.20	0.14	0.39	0.23	0.16	0.16	0.43	0.96	0.95

Note: Bootstrap standard errors are given in parentheses.

(b) Precinct-specific estimates of β_i^b and β_i^w for the Carter data set

Precinct	β_i^b				β_i^w			
	King	Method I	Method II	Method III	King	Method I	Method II	Method III
1	0.9580	0.9783	0.9540	0.9556	**0.1672**	**0.1642**	**0.1678**	**0.1675**
2	**0.9683**	**0.9643**	**0.9503**	**0.9588**	0.1250	0.1331	0.1616	0.1443
3	0.9383	0.9587	0.8484	0.9927	**0.0591**	**0.0555**	**0.0765**	**0.0490**
4	0.9500	0.9730	0.8914	0.9786	**0.1063**	**0.1053**	**0.1087**	**0.1051**
5	0.9546	0.9602	0.9108	0.9701	**0.1119**	**0.1092**	**0.1335**	**0.1043**
6	0.9592	0.9760	1.0000	0.9346	**0.2151**	**0.2148**	**0.2143**	**0.2156**
7	0.9474	0.9739	0.9121	0.9703	**0.1272**	**0.1261**	**0.1287**	**0.1262**
8	0.9562	0.9745	0.9286	0.9635	**0.1438**	**0.1434**	**0.1443**	**0.1436**
9	0.9700	0.9921	0.9775	0.9542	**0.1874**	**0.1775**	**0.1840**	**0.1945**
10	**0.9683**	**0.9683**	**0.9680**	**0.9683**	0.1428	0.1380	0.1695	0.1416
11	**0.9924**	**0.9927**	**0.9905**	**0.9930**	0.1448	0.1391	0.1813	0.1322

Note: Estimates of β_i^b for majority black precincts and of β_i^w for majority white districts are indicated in bold.

see this consistency, we turn to working through two real-world examples: an 11-precinct sample of data (Carter data set) from a biracial legislative contest in a Deep South state in the 1980s, and a 284-precinct sample from the 1995 gubernatorial race in Louisiana, which included a prominent black candidate, Cleo Fields.[11] Our first data set was chosen to have a small number of precincts so as to demonstrate that even for small data sets, if patterns are clear enough, it is relatively easy to see what is happening.

For the Carter data set, King's EZI gets a mean value of (B^b, B^w) of (.978, .136) (see Table 5.3a). Our three estimates for the district parameters are (.974, .138), (.940, .155), and (.964, .144) for Methods I, II, and III, respectively. Note that, even though Method II has the potential to give results quite different from Methods I and III or from King's basic EZI program, in this real-world data set the three methods give results that are not very different. Furthermore,

[11] Cleo Fields was a state senator and former U.S. representative from majority black districts (see Voss, 1999, for further background).

Table 5.4 Comparison of Method I estimates, unweighted and weighted for informativeness

Data set	Parameter	Method I unweighted	Method I weighted	King's method
			Value	
Data set C	B^b	.829	.957	.917
	B^w	.171	.043	.083
Carter 11-	B^b	.974	.972	.978
Precinct set	B^w	.138	.139	.136
Cleo Fields	B^b	.874	.900	.901
284-precinct	B^w	.025	.015	.014
set				

for the Carter data set (see Table 5.4) the Method I estimates weighted for informativeness are (.972, .139), almost identical with the unweighted Method I estimates (.974, .138).

We can also compare precinct-specific values of β_i^b and β_i^w. We show those estimates for our three methods, along with the corresponding values from King's truncated normal MLE method, in Table 5.3b. Again there is very high consistency in the estimates, particularly – as expected – for estimates of β_i^b for majority black precincts and of β_i^w for majority white districts. These precincts are identified in bold in Table 5.3b. Indeed, the maximum difference between our Method I and King's method in estimating the β_i^b and β_i^w values for the majority race is 0.004, while the maximum difference for the corresponding estimates for the minority race is 0.027. These results are summarized in Figure 5.5, which plots the precinct-level estimates β_i^b (Figure 5.5a) and β_i^w (Figure 5.5b) for each of our three methods and King's method versus the black proportion of the population. It is apparent that – except for Method II – the estimates of β_i^b are almost identical to each other for heavily black precincts and that the estimates of β_i^w are almost identical to each other for heavily white precincts.

The Cleo Fields data set involves 284 precincts, many of which were heavily white or heavily black. We would expect these to be most informative about the voting behavior of the respective majority races. Table 5.4 compares the parameter estimates for Method I – weighted and unweighted – with those of King, for both the Carter and Cleo Fields data sets as well as for data set C. Unlike the Carter data set, the Cleo Fields data set shows significant differences when weighting is used; the weighted estimates of Method I are almost identical with those of King.

5.6 DISCUSSION

Since the least-squares and area minimization approaches we have offered are less general than King's method,[12] why might anyone care about results derived from them? There are several reasons. First, each generates closed-form solutions that are trivial to calculate, albeit the distribution of precinct specific values each generates cannot be characterized as a particular standard type (e.g., a truncated bivariate normal). In particular, because of the simplicity of the numerical calculations needed for our methods, it is practical to extend

---

[12] For example, to be comparable to the more advanced versions of King's program we would need to develop some explicit way of building in covariates.
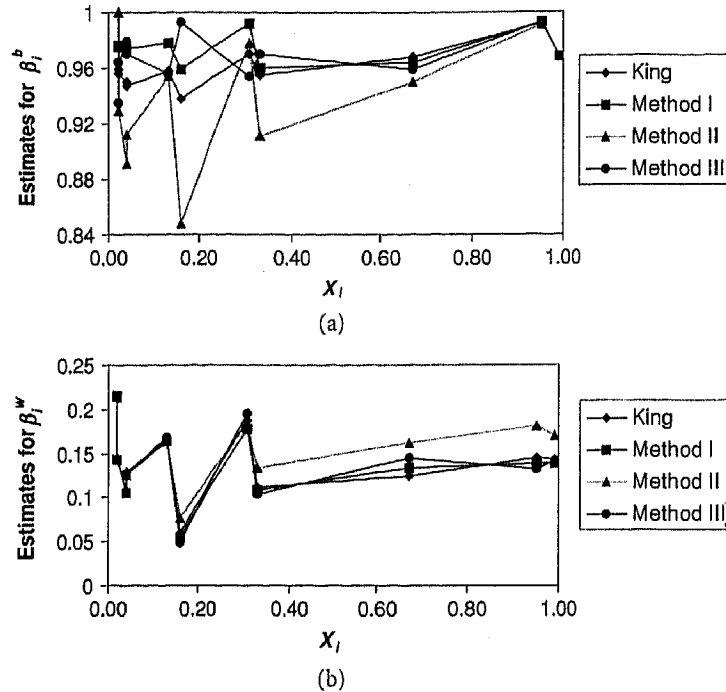
Figure 5.5. Precinct-level Estimates of (a) β_i^b and (b) β_i^w by King's method and our three methods for Carter data.

them from 2 × 2 to m × n tables, as has been shown by De Sio (2003). Second, each has a very simple mathematical exposition,[13] and can be described in reasonably intuitive terms.[14] Third, each offers either an indirect or a direct analogue to the OLS approach to ecological regression, but with the advantage that parameter results are constrained to feasible values. Fourth, although the following question needs considerable more investigation, based on our explorations so far, it appears likely that two of these simple methods (Methods I and III) will closely approximate the results generated by the basic MLE approach in King's (bivariate) EZI program. Finally, a comparison of the simple optimization aspects of our methods with the MLE approach of King, and comparisons among the three methods we introduce, shed some light on how ecological inference works and when it might be expected to fail.

APPENDIX 1. PROOF OF THEOREM 1

Theorem 1. If the tomographic line segment bounds used as the basis for King's ecological inference have a unique intersection, this intersection will be at a point, $(\hat{B}^b, \hat{B}^w)$, which corresponds to the (B^b, B^w) values derived from Goodman's method of bivariate ecological regression.

[13] The Monte Carlo nature of the MLE estimation procedure, combined with the complexities of estimating a truncated normal distribution, renders King's procedures much more complex than those we propose.

[14] Grofman and Merrill (2002) discuss criteria for a good solution to the problem of ecological inference, and Silva de Mattos and Veiga (this volume, Chapter 15) assess the predictive ability of several methods, including those of Goodman and King.

Proof. If $(\hat{B}^b, \hat{B}^w)$ lies on all the constraint lines of a tomographic plot, the coordinates $\hat{B}^b$ and $\hat{B}^w$ must satisfy

$$T_i = \hat{B}^b X_i + (1 - X_i)\hat{B}^w$$

for all i. Thus, after averaging, we have

$$T = \hat{B}^b X + (1 - X)\hat{B}^w$$

as well, where X and T are the (weighted) averages for the entire district. But the slope coefficient of ecological regression is given by

$$\hat{m} = \frac{\sum (X_i - X) T_i}{\sum (X_i - X)^2},$$

where

$$\begin{aligned}
\sum (X_i - X) T_i &= \sum (X_i - X)\left[\hat{B}^b X_i + (1 - X_i)\hat{B}^w\right] \\
&= \sum X_i - X)\left[X_i(\hat{B}^b - \hat{B}^w) + \hat{B}^w\right] \\
&= \sum (X_i - X)\left[(X_i - X)(\hat{B}^b - \hat{B}^w)\right] + \sum (X_i - X)\hat{B}^w \\
&= \sum (X_i - X)^2\left[\hat{B}^b - \hat{B}^w\right],
\end{aligned}$$

and where we have twice used the fact that $\sum (X_i - X) = 0$.

It follows that

$$\hat{m} = \hat{B}^b - \hat{B}^w.$$

But then the intercept of ecological regression is given by

$$\begin{aligned}
\hat{b} &= T - \hat{m}X = (\hat{B}^b X + (1 - X)\hat{B}^w) - (\hat{B}^b X - \hat{B}^w X) \\
&= \hat{B}^w,
\end{aligned}$$

so that $\hat{B}^b$ and $\hat{B}^w$ are equal to the parameter estimates of ecological regression. *q.e.d.*

APPENDIX 2. AREA CALCULATIONS FOR METHOD III

Before we tackle the general problem of minimizing the areas between the line $m = (T - b)/X$ and the bounding line segments of the cones associated with the precinct values X_i and T_i, we begin by specifying the area between any two lines contained in the range $X \in [0, 1]$. Let one line be $T_{(1)} = m_1 X + b_1$, and the other be $T_{(2)} = m_2 X + b_2$. These lines intersect at $X_c = -(b_2 - b_1)/(m_2 - m_1) = -\Delta b/\Delta m$, where $\Delta b = b_2 - b_1$ and $\Delta m = m_2 - m_1$.

There are two cases, depending on whether $X_c \in [0, 1]$, i.e., whether $0 \le -\Delta b/\Delta m \le 1$. We summarize cases and results in the table below:

Case	Intersection	Area between two lines
I	Outside $X_c \in [0, 1]$	$\left\| \dfrac{\Delta m}{2} + \Delta b \right\|$
II	Within $X_c \in [0, 1]$	$\left\| \dfrac{(\Delta b)^2}{\Delta m} + \dfrac{\Delta m}{2} + \Delta b \right\|$

To verify the results given in the table, we look at the relevant integrals. For case I, the area between the two lines is given by the absolute value of

$$\int_0^1 \left[T_{(2)} - T_{(1)} \right] dX = \int_0^1 \left[m_2 X + b_2 - m_1 X - b_1 \right] dX = \frac{1}{2} \Delta m + \Delta b.$$

Similarly, for case II, if $b_1 < b_2$, the areas between the two lines are given by

$$\int_0^{-\Delta b/\Delta m} \left[T_{(2)} - T_{(1)} \right] dX + \int_{-\Delta b/\Delta m}^1 \left[T_{(1)} - T_{(2)} \right] dX$$

$$= \int_0^{-\Delta b/\Delta m} \left[m_2 X + b_2 - m_1 X - b_1 \right] dX + \int_{-\Delta b/\Delta m}^1 \left[m_1 X + b_1 - m_2 X - b_2 \right] dX$$

$$= \int_0^{-\Delta b/\Delta m} \left[(\Delta m) X + \Delta b \right] dX - \int_{-\Delta b/\Delta m}^1 \left[(\Delta m) X + \Delta b \right] dX$$

$$= \frac{1}{2} \frac{(\Delta b)^2}{\Delta m} - \frac{(\Delta b)^2}{\Delta m} + \frac{1}{2} \frac{(\Delta b)^2}{\Delta m} - \frac{(\Delta b)^2}{\Delta m} - \frac{1}{2} \Delta m - \Delta b$$

$$= -\frac{(\Delta b)^2}{\Delta m} - \frac{1}{2} \Delta m - \Delta b.$$

If, instead, $b_1 \geq b_2$, the same result is obtained without the negative signs. In each case, the area is positive and hence is the absolute value of the expression. q.e.d.

Now, we can solve the general problem by finding the line whose b value minimizes the sum of the appropriate areas for each pairing between that line and the other lines in the set, subject to the constraint that the m and b values of that line must satisfy the equation $T = mX + b$.

REFERENCES

Anselin, Luc and Wendy Tam Cho. 2002. "Spatial Effects and Ecological Inference." *Political Analysis*, 10: 276–297.

Cho, Wendy Tam. 1998. "Iff the Assumption Fits ... : A Comment on the King Ecological Inference Solution." *Political Analysis*, 7: 143–163.

Cho, Wendy Tam. 2001. "Latent Groups and Cross-Level Inferences." *Electoral Studies*, 20, 2: 243–263.

Cho, Wendy Tam and Albert H. Yoon. 2001. "Strange Bedfellows: Politics, Courts, and Statistics: Statistical Expert Testimony in Voting Rights Cases." *Cornell Journal of Law and Public Policy* 10, 2: 237–264.

De Sio, Lorenzo. 2003. "A Proposal for Extending King's EI Method to $m \times n$ Tables." Transcript: University of Florence.

Duncan, Dudley and Beverley Davis. 1953. "An Alternative to Ecological Correlation." *American Sociological Review*, 18: 665–666.

Efron, Bradley and Robert Tibshirani. 1993. *An Introduction to the Bootstrap*. London: Chapman and Hall.

Freedman, David, S.P. Klein, J. Sacks, C.A. Smyth and C.G. Everett. 1991. "Ecological Regression and Voting Rights." *Evaluation Review*, 15, 6: 673–711.

Freedman, David, M. Ostland, M.R. Roberts, and S.P. Klein. 1999. "The Future of Ecological Inference Research: A Comment on Freedman et al. – Response to King's Comment. *Journal of the American Statistical Association*, 94, 445: 355–357.

Goodman, Leo. 1953. "Ecological Regression and the Behavior of Individuals." *American Sociological Review*, 18, 6: 663–664.

Goodman, Leo. 1959. "Some Alternatives to Ecological Correlation." *American Journal of Sociology*, 64: 610–625.

Grofman, Bernard. 1991. "Statistics without Substance: A Critique of Freedman et al. and of Clark and Morrison." *Evaluation Review*, 125, 6: 746–769.

Grofman, Bernard. 1992. "Expert Witness Testimony and the Evolution of Voting Rights Case Law." In Bernard Grofman and Chandler Davidson (eds.), *Controversies in Minority Voting: The Voting Rights Act in Perspective*. Washington, DC Brookings Institution.

Grofman, Bernard. 1993. "Voting Rights in a Multi-ethnic World." *Chicano-Latino Law Review*, 13, 15: 15–37.

Grofman, Bernard. 2000. "A Primer of Racial Bloc Voting." In Nathaniel Persily (ed.), *The Real Y2K Problem: Census 2000 Data and Redistricting Technology*. Brennan Center for Law and Justice, New York University, pp. 44–67.

Grofman, Bernard and Samuel Merrill, III. 2002. "What Does It Mean to Offer a 'Solution' to the Problem of Ecological Inference?" Typescript: University of California–Irvine.

Grofman, Bernard and Michael Migalski. 1988. "Estimating the Extent of Racially Polarized Voting in Multicandidate Elections." *Sociological Methods and Research*, 16, 4: 427–454.

King, Gary. 1997. *A Solution to the Ecological Inference Problem*. Princeton, NJ: Princeton University Press.

King, Gary, Ori Rosen, and Martin Tanner. 1999. "Binomial-Beta Hierarchical Models for Ecological Inference." *Sociological Methods and Research*, 28, 1: 61–90.

McCue, Kenneth. 2001. "The Statistical Foundations of the EI Method." *American Statistician*, 55, 2: 106–110.

Voss, D. Stephen. 1999. "Racial Redistricting and the Quest for Legislative Diversity." Extensions of Remarks. *APSA Legislative Studies Section Newsletter* 22, July:11–14.