Logical Inconsistency in King-based Ecological Regressions[1]

Michael C. Herron[2]
Department of Political Science
Northwestern University

Kenneth W. Shotts[3]
Department of Political Science
Northwestern University

July 3, 2002

[1]An earlier draft of this paper was presented at the 2000 Annual Meeting of the American Political Science Association. The authors thank Wendy Tam Cho and Jasjeet Sekhon for comments.
[2]601 University Place, Evanston, IL 60208–1006 (m–herron@northwestern.edu).
[3]601 University Place, Evanston, IL 60208–1006 (k–shotts@northwestern.edu).

Abstract

The statistical procedure EI–R, in which point estimates produced by the King (1997) ecological inference technique are used as dependent variables in a linear regression, can be logically inconsistent insofar as the assumptions necessary to support EI–R's first stage (ecological inference via King's method) can be incompatible with the assumptions supporting its second stage (linear regression). In light of this problem, we derive a specification test for logical consistency of EI–R and describe options available to a researcher who confronts test rejection. We then apply our test to the implementation of EI–R in Burden and Kimball's (1998) study of ticket splitting and find that this implementation is logically inconsistent. In correcting for this problem we show that Burden and Kimball's alleged substantive results are not results at all and instead are artifacts of a self-contradictory statistical technique.

1 Introduction

It has become quite common for researchers to use point estimates produced by the King (1997) ecological inference technique as dependent variables in second stage linear regressions. This two-stage statistical procedure, which Herron and Shotts (2001) call EI–R, is strongly advocated by King, who says that "[EI–R] has enormous potential for uncovering new information" (p. 279). Burden and Kimball (1998), for example, use King–based estimates of district–level ticket splitting rates as dependent variables in second stage regressions which examine why individuals cast split ticket ballots. Similarly, Gay (2001) employs estimates of black turnout rates, produced by King's ecological inference technique, in regressions that seek to determine whether black citizens vote at higher rates in elections that include black Congressional candidates. Other examples of EI–R can be found in Kimball and Burden (1998), Voss and Lublin (1998), Cohen, Kousser and Sides (1999), Wolbrecht and Corder (1999), Voss and Miller (2000), Burden and Kimball (2002), and Lublin and Voss (2002).

Clearly, EI–R has achieved widespread usage despite the fact that it was only recently proposed. We show, however, that applications of this two-stage procedure can be *logically inconsistent* insofar as the assumptions necessary to support the procedure's first stage (ecological inference via King's method) can be incompatible with the assumptions supporting its second stage (linear regression). In contrast, we say that an application of EI–R is *logically consistent* when its first stage and second stage assumptions are not mutually contradictory.

The matter of EI–R's being logically consistent merits attention because EI–R has two stages, unlike, say, a regular regression that has only a single stage. Whenever a researcher employs a two-stage statistical procedure, she must simultaneously adopt assumptions that support the procedure's first and second stages. Previous work has not sought to determine whether this is possible in the case if EI–R. Indeed, to the best of our knowledge all published and working–paper implementations of EI–R simply assume that the procedure's two stages are inherently compatible.

Such a blithe attitude is not warranted, and we show here that the assumptions necessary to support EI–R can be self–contradictory. This result follows from the fact that standard implementations of King's technique to an ecological dataset (the first stage of EI–R) assume that there is no aggregation bias in the dataset. Nonetheless, a linear regression which uses King–based point estimates as dependent variables (the second stage of EI–R) can imply the existence of aggregation bias.

To illustrate this clash of assumptions, suppose that a researcher wishes to estimate and explain disaggregated turnout rates of African-American and white voters but for this task has access only to precinct–level racial composition and turnout data. To apply King's ecological inference technique to her turnout data and estimate both African-American and white turnout rates by precinct (first stage of EI–R), the researcher must assume that a precinct's black turnout rate is not a function of the extent to which it contains African-Americans, i.e., there is no aggregation bias in her data.

Now, suppose that the researcher estimates a regression (second stage of EI–R) in which black turnout rates are modeled as a linear function of precinct–level income and education data. If the regression's right hand side covariates are correlated with the fraction of a precinct's population that is African-American, then the researcher's second stage regression virtually guarantees (in a way we make precise later) that a precinct's black turnout rate is a function of the fraction of its population that is African-American, i.e., there is aggregation bias. Thus, in this example the assumptions behind the second stage of EI–R contradict the assumptions in its first stage.

This is problematic at two levels. At a fundamental level, any standard philosophy of science requires researchers to adopt internally consistent sets of assumptions. To do otherwise, as is possible with EI–R, is to state simultaneously "**A** is true" and "**A** is false." Once two mutually contradictory premises are adopted, any implication can be logically derived, i.e., all subsequent analysis is meaningless.

At a more pragmatic level, logical inconsistency of EI–R is problematic because EI–R, like all statistical techniques, must ultimately depend on a set of assumptions, often called regularity conditions. Although King (1997) does not present a set of sufficient conditions under which EI–R estimates are well-behaved, i.e., consistent, asymptotically normal, and so forth, any such set must consist only of assumptions that are not self-contradictory. Therefore, when EI–R is logically inconsistent, the estimates it produces have no known properties. No one knows, that is, if the estimates produced by a logically inconsistent application of EI–R are unbiased, consistent, and so forth. This problem is independent of whether a software implementation of EI–R produces "standard errors" for EI–R estimates. Standard errors only have meaning if they are grounded in a logically consistent set of assumptions.

Given the problems that result if EI–R is logically inconsistent, it is important for researchers to determine whether a particular application of the technique is problematic in this way. In theory—and fortuitously, one might argue—not all implementations of EI–R are logically inconsistent. Indeed, we show that logical consistency of EI–R is application–dependent insofar as EI–R is logically consistent in some empirical applications yet logically inconsistent in others (although, as we explain later, it is exceedingly hard to imagine that politically interesting applications of EI–R are logically consistent).

Thus, we derive a specification test for EI–R, and the test is implemented with a series of bivariate regressions. Interpreting the output of our test is simple: if any of the slope estimate t–statistics produced by the bivariate regressions are statistically significant (at a pre-specified confidence level), then EI–R as applied to the dataset is logically inconsistent and its estimates have no known properties. If, on the other hand, all of the t values are insignificant, then logical consistency of EI–R cannot be rejected.

When the test for logical consistency of EI–R rejects in a given ecological dataset, a researcher who wants to study the dataset has only two options. One, she can abandon EI–R entirely and choose a different method of data analysis. Or, two, she can explicitly incorporate the covariates that were to be used in a second stage regression into the first stage of the EI–R procedure and use what King calls the extended ecological inference model. This model, which we discuss in detail later, has only one stage and therefore is not at risk of logical inconsistency. In addition, King's extended model is designed to handle aggregation bias in ecological data, and this is useful since the EI–R logical inconsistency on which we focus is caused by unmodeled aggregation bias.

Several things about the relationship of our analysis to previous research are worth noting. First, our recommendation of King's extended model as a replacement for EI–R when the latter is logically inconsistent needs to be understood in light of the difficulty of ecological inference (e.g., Achen and Shively 1995). Indeed, no one should engage in ecological inference without recognizing that the ecological inference problem cannot be solved in any meaningful way.[1] Rather, all proposed solutions to this problem are based on strong and often untestable assumptions about aggregate data.

Second, our critique of EI–R and King's ecological inference technique differs fundamentally from the arguments in Tam (1998) and Cho and Gaines (2000), both of which claim that aggregation bias is a common phenomenon in ecological data and that diagnostics advocated by King (1997, ch. 9) provide little help in identifying and rectifying it. We do not claim that our test for logical consistency of EI–R is a general test for aggregation bias. Rather, our critique applies only to the use of King's technique within the context of the combined EI–R procedure, and what we argue is that the assumptions which support EI–R must be compatible across EI–R's two stages.

In Section 2 we provide notation and details on EI–R. Section 3 then presents our specification test for logical consistency of EI–R and explains how to interpret it. Section 4 considers the options faced by a researcher whose ecological dataset fails the test, and in Section 5 we apply the test to a dataset from Burden and Kimball's (1998) study of ticket splitting. This dataset easily fails the specification test, and thus we use King's extended model to analyze it. The results of our extended model analysis are dramatically different from Burden and Kimball's published findings on the causes of ticket splitting, and this illustrates the costs of ignoring logical inconsistency in EI–R. Section 6 offers general comments on EI–R and concludes.

[1] This is because, simply put, an ecological inference problem consists of a system of p equations with $2p$ unknowns. There is no unique solution to such a system.

2 Overview of Ecological Inference and EI–R

This section provides some details on King's ecological inference technique and explains how the technique's output is used in second stage regressions.

2.1 Initial Details on King's Ecological Inference Technique

Suppose that we have precinct–level data for an election: let $T_i \in [0,1]$, $i = 1, \ldots, p$, denote precinct i's turnout rate in the election and $X_i \in [0,1]$ the fraction African-American of precinct i. The goal of ecological inference is using T_i and X_i to estimate β_i^b, the fraction of blacks in precinct i who voted in the election, and β_i^w, the fraction of whites who voted.

King develops two different implementations of his ecological inference technique, *standard* and *extended*. Details of these implementations are in King (1997, ch. 6–8). Critiques of King's approach to ecological inference can be found in Freedman, Klein, Ostland and Roberts (1998), Tam (1998), Cho and Gaines (2000), and McCue (2001). Until further notice our comments on King's technique apply to the standard model.

King's standard model is grounded in the accounting identity for T_i:

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

Equation (1) reflects the fact that total turnout in precinct i is the precinct's black turnout plus its white turnout. King's model assumes that $\left(\beta_i^b, \beta_i^w\right)$ have a truncated bivariate normal distribution on the unit square and that β_i^b (β_i^w) can be decomposed into a mean $\mathfrak{B}^b$ ($\mathfrak{B}^w$) and a deviation from this mean ϵ_i^b (ϵ_i^w). It therefore follows from equation (1) that

$$\begin{aligned} T_i &= \left(\mathfrak{B}^b + \epsilon_i^b\right) X_i + \left(\mathfrak{B}^w + \epsilon_i^w\right) (1 - X_i) & (2) \\ &= \mathfrak{B}^b \, X_i + \mathfrak{B}^w \, (1 - X_i) + \epsilon_i^b \, X_i + \epsilon_i^w \, (1 - X_i) & (3) \\ &= \mathfrak{B}^w + X_i \left(\mathfrak{B}^b - \mathfrak{B}^w\right) + \epsilon_i, & (4) \end{aligned}$$

where ϵ_i^b and ϵ_i^w are mean zero disturbances and where $\epsilon_i = X_i \, \epsilon_i^b + (1 - X_i) \, \epsilon_i^w$. The means

$(\mathfrak{B}^b, \mathfrak{B}^w)$, standard deviations (σ_b, σ_w), and correlation (ρ) of β_i^b and β_i^w are collected in a 5–vector $\psi = \{\mathfrak{B}^b, \mathfrak{B}^w, \sigma_b, \sigma_w, \rho\}$.[2]

Three facts about the above equations are worth noting. First, ϵ_i^b in equation (2) reflects unmodeled variation in β_i^b just as a disturbance in a regular linear regression reflects unmodeled terms that affect the regression's dependent variable (and similarly for ϵ_i^w and β_i^w). Second, since it follows from an accounting identity, equation (4) holds exactly. Third, King's technique requires that the disturbance ϵ_i, conditional on X_i, be mean zero. Indeed, since King assumes $E\left(\epsilon_i^b|X_i\right) = E\left(\epsilon_i^w|X_i\right) = 0$, where $E(\cdot)$ denotes expectation, ϵ_i has this property.[3]

These twin assumptions on the black and white disturbances, ϵ_i^b and ϵ_i^w, respectively, imply that there is no aggregation bias. Substantively speaking, if $E\left(\epsilon_i^b|X_i\right) = 0$ then, once fraction black X_i in precinct i is known, there is nothing in the random disturbance ϵ_i^b that affects β_i^b and is also mean dependent on X_i (and similarly for ϵ_i^w and β_i^w). It cannot be the case, according to King's standard ecological inference technique, that ϵ_i^b includes a non-random component (e.g., median family income in precinct i) that is a linear function of X_i (percent black in precinct i) and also influences β_i^b (black turnout in precinct i).

Let $\hat{\mathfrak{B}}^b$, $\hat{\mathfrak{B}}^w$, $\hat{\sigma}_b$, $\hat{\sigma}_w$, and $\hat{\rho}$ denote the point estimates of $\mathfrak{B}^b$, $\mathfrak{B}^w$, σ_b, σ_w, and ρ, respectively, based on applying King's standard model to an ecological dataset.[4] These five point estimates are collected in a 5–vector $\hat{\psi}$. Assuming that the standard model's regularity conditions hold, $\hat{\mathfrak{B}}^b \xrightarrow{\text{Pr}} \mathfrak{B}^b$ where $\xrightarrow{\text{Pr}}$ denotes convergence in probability. In other words, $\hat{\mathfrak{B}}^b$ is a consistent estimate of $\mathfrak{B}^b$, and a similar statement applies to the other four point estimates in $\hat{\psi}$. In general, then, $\hat{\psi} \xrightarrow{\text{Pr}} \psi$. Let $\hat{\beta}_i^b$ $(\hat{\beta}_i^w)$ be the point estimate of β_i^b (β_i^w) produced by applying King's technique to an ecological dataset X_i and T_i, $i = 1, \ldots, p$. See McCue (2001) for a discussion of these estimates, which by construction are functions of $\hat{\psi}$, X_i, and T_i.

[2]The elements of ψ correspond to elements in a related 5–vector $\breve{\psi}$ where ψ is a function of $\breve{\psi}$. We ignore $\breve{\psi}$ because the distinction between it and ψ is numerical rather than substantive.

[3]In addition, King's technique assumes that there is no spatial autocorrelation, i.e., that T_i and T_j are independent for $i \neq j$.

[4]King's standard model can be implemented using likelihood theory or priors can be employed in which case the model is Bayesian. Both implementations are compatible our analysis.

2.2 Definition of EI–R

In EI–R, $\hat{\beta}_i^b$ (or $\hat{\beta}_i^w$, as discussed below) point estimates are used as dependent variables in a second stage regression. Namely, King suggests that a researcher who wants to estimate

$$\beta_i^b = \gamma_0 + \gamma' Z_i + \nu_i^b, \tag{5}$$

instead estimate

$$\hat{\beta}_i^b = \gamma_0 + \gamma' Z_i + \nu_i^b, \tag{6}$$

where the difference between equations (5) and (6) is that $\hat{\beta}_i^b$ (observed) substitutes for β_i^b (unobserved). Here, γ_0 is a constant, γ a k–vector of slope parameters to be estimated, Z_i a k–vector of covariates, and ν_i^b a disturbance assumed to be uncorrelated with $\hat{\beta}_i^b$. Let $\hat{\gamma}_0$ and $\hat{\gamma}$ denote the least squares estimates of γ_0 and γ, respectively.

Thus, EI–R consists of, first, estimating $\hat{\beta}_i^b$ using King's ecological inference technique and, second, estimating equation (6) with least squares.[5] The final products of EI–R, therefore, are the estimates $\hat{\gamma}_0$ and $\hat{\gamma}$, and, ultimately, whether EI–R is statistically consistent depends on whether $\hat{\gamma}_0 \xrightarrow{\text{Pr}} \gamma_0$ and $\hat{\gamma} \xrightarrow{\text{Pr}} \gamma$. On the other hand, first stage statistical consistency of EI–R depends on whether $\hat{\psi} \xrightarrow{\text{Pr}} \psi$ where the 5–vector $\hat{\psi}$, as noted previously, is produced by applying King's standard model to an ecological dataset.

Of course, equations (5) and (6) could just as easily have had β_i^w and $\hat{\beta}_i^w$, respectively, on their left hand sides. In this case, EI–R is based on

$$\beta_i^w = \alpha_0 + \alpha' Z_i + \nu_i^w, \tag{7}$$

even though in practice it estimates

$$\hat{\beta}_i^w = \alpha_0 + \alpha' Z_i + \nu_i^w. \tag{8}$$

[5] King suggests using weighted least squares when estimating equation (6). Weighting, however, is not germane to our discussion of logical inconsistency.

There is no loss of generality in assuming that the same covariate vector Z_i appears in both equations (5) and (7), and there is also no loss of generality in assuming that Z_i is mean zero.

3 The Logical Consistency Test for EI–R

The two underlying regressions in a given application of EI–R are equations (5) and (7). Substituting these into the accounting identity of equation (1) yields

$$\begin{aligned} T_i &= \alpha_0 + X_i\ (\gamma_0 - \alpha_0) + X_i\ (\gamma - \alpha)' Z_i + \alpha' Z_i + \nu_i^b X_i + \nu_i^w\ (1 - X_i) && (9) \\ &= \alpha_0 + X_i\ (\gamma_0 - \alpha_0) + \tilde{\epsilon}_i, && (10) \end{aligned}$$

where $\tilde{\epsilon}_i = X_i\ (\gamma - \alpha)' Z_i + \alpha' Z_i + \nu_i^b X_i + \nu_i^w\ (1 - X_i)$. To be logically consistent, a researcher who uses EI–R must simultaneously adopt the assumptions that support both EI–R's first and second stages. Therefore, it has to be true that $E(\tilde{\epsilon}_i|X_i) = 0$ since, with reference to the presence of ϵ_i in equation (4), this is precisely what was assumed earlier.

We henceforth assume $E\left(\nu_i^b|X_i\right) = E\left(\nu_i^w|X_i\right) = 0$. This conservative assumption guarantees that two of the four components of $\tilde{\epsilon}_i$ are mean zero conditional on X_i, as required for logical consistency of EI–R. However, this is not enough insofar as our assumptions about ν_i^b and ν_i^w do not ensure that $E(\tilde{\epsilon}_i|X_i) = 0$. Rather, for this latter equality to hold it must also be true that $E\left(X_i\ (\gamma - \alpha)' Z_i + \alpha' Z_i|X_i\right) = 0$, where $E\left(X_i\ (\gamma - \alpha)' Z_i + \alpha' Z_i|X_i\right) = (X_i(\gamma - \alpha) + \alpha)' E(Z_i|X_i)$. We have, therefore, just derived the following result:

Proposition 1 *For EI–R to be logically consistent, it must be true that* $(X_i(\gamma - \alpha) + \alpha)' E(Z_i|X_i) = 0$.

If $\gamma = \alpha = 0$ then the restriction in Proposition 1 clearly holds. This is a trivial case, of course, since one would never want to assume $\gamma = \alpha = 0$. Doing so assumes that all second stage slope coefficients are zero and that there is nothing to estimate! We cannot, therefore, assume that the restriction in Proposition 1 holds on account of $\gamma = \alpha = 0$, and we must look elsewhere to see what might help with it.

Suppose that $E(Z_i|X_i) \neq 0$. Is it possible in this circumstance that the restriction in Proposition 1 holds? Unless the slope vectors γ and α satisfy an extremely tight and arbitrary restriction, the answer to this question is no. To see this, suppose for purposes of exposition that α and γ are scalars. Then, $E(Z_i|X_i) \neq 0$ implies that the restriction in Proposition 1 holds only if $X_i(\gamma - \alpha) + \alpha = 0$, which simplifies to $\gamma = \alpha(1 - 1/X_i)$. It is obvious that, if X_i takes on more than two different strictly positive values, which it must for the first stage of EI–R to be identified, then there are no values of γ and α that are consistent with $\gamma = \alpha(1 - 1/X_i)$ except $\gamma = \alpha = 0$.

This logic is easily generalizable to the case where α and γ are vectors. In this general case, if $E(Z_i|X_i) \neq 0$, then the restriction in Proposition 1 holds only under extremely tight constraints on γ and α. These latter constraints are non-generic meaning that, if γ and α were drawn from a continuous probability distribution, they would hold with probability zero.

On the other hand, if $E(Z_i|X_i) = 0$, then the restrictions in Proposition 1 hold regardless of γ and α. Thus,

Proposition 2 *Barring trivial and probability zero circumstances, logical consistency of EI–R requires* $E(Z_i|X_i) = 0$.

Recall that Z_i was assumed, without loss of generality, to be mean zero. Therefore, Proposition 2 requires that the unconditional expectation of Z_i be identical to its expectation conditional on X_i.

If, however, X_i does contain information about the mean of Z_i so that $E(Z_i|X_i) \neq 0$, then it follows that, barring trivial and probability zero circumstances, $E(\tilde{\epsilon}_i|X_i) \neq 0$. Since $\tilde{\epsilon}_i = \epsilon_i^b + \epsilon_i^w$, $E(Z_i|X_i) \neq 0$ therefore implies that either ϵ_i^b and/or ϵ_i^w from EI–R's first stage contains unmodeled variables. Thus, when the restriction in Proposition 2 does not hold, then the second stage of EI–R implies that its first stage is affected by aggregation bias.

It is straightforward to test whether, in a given application of EI–R, the conditional expectations in Proposition 2 are zero. But before we explain how to do this, it is useful to think in a substantive way about what Proposition 2 implies.

Consider Gay's (2001) EI–R analysis of the precinct-level socioeconomic covariates that affect black and white turnout. In Gay's framework, as in the running example that has been used here, X_i is the fraction African-American of precinct i, β_i^b (β_i^w) is precinct i's black (white) turnout rate, and Z_i is a vector of covariates assumed to affect β_i^b and β_i^w.

It is quite reasonable to believe that elements of Gay's Z_i vector, e.g., precinct-level education and income levels in her Congressional race analyses of Michigan and Pennsylvania, do in fact influence voter turnout levels (e.g., Mebane and Sekhon 2002). We would therefore expect a regression of β_i^b on Z_i to return significant results. But, of course, β_i^b cannot be observed and this leads Gay to EI–R which calls for estimating a regression of $\hat{\beta}_i^b$ on Z_i.[6]

For Gay's Michigan and Pennsylvania EI–R analyses to be logically consistent, it must be true that knowledge of a precinct's percent black (X_i) tells us nothing about the precinct's per capita income (an element of Gay's Z_i). This is untenable in light of contemporary American social realities: if a precinct has a large African-American population, then all things equal this precinct will have a relatively low per capita income. Nonetheless, without assuming that a precinct's per capita income is not a function of its racial composition, and without making a host of similarly implausible assumptions for the other right hand side variables in her second stage regressions, Gay's use of EI–R is logically inconsistent.

In general, then, the mathematical restrictions in Proposition 2 have very strong substantive implications. Indeed, it is hard to imagine any politically interesting applications of EI–R that contain Z_i vectors unrelated to X_i levels as required in the proposition. In theory, however, such applications might exist, and hence it is important to develop a specification test for logical consistency of EI–R.

To test the conditional expectation restrictions in Proposition 2 on a given ecological dataset, one regresses each element of Z_i—recall that Z_i is a k vector—on X_i. If any of the resulting k slope estimates are statistically significant (intercepts are irrelevant, since as

[6]Gay (2001) estimates many second stage regressions; see her Tables 3 and 4. Thus, our reference to "Gay's EI–R analysis" is a generic way of discussing any of these regressions. We focus on Gay's Michigan and Pennsylvania analyses in the text because these include socioeconomic variables in a second stage Z_i vector. Nonetheless, Gay's other analyses almost certainly suffer from logical inconsistency as well.

noted before we can assume without loss of generality that the mean of Z_i is zero) then it follows that, barring trivial and probability zero circumstances, the restriction in Proposition 2 does not hold. Thus, our test for logical consistency of EI–R requires nothing more than estimating linear regressions with data that EI–R itself requires in the first place.

It is important to recognize that the test we propose for logical consistency of EI–R is one of necessity as opposed to sufficiency. That is, an application of EI–R must pass our test; if not, it is logically inconsistent. However, if an application of EI–R does pass the test, it does not therefore follow that the application is logically consistent. Rather, as we discuss below, it is still possible that the application is logically inconsistent in some manner not covered by our test. Or, the failure to reject could simply reflect that fact that our test does not have a power of one.

4 What to do when an Ecological Dataset Fails the Specification Test

When a researcher's ecological dataset fails the test for logical consistency of EI–R, he has two options. One, he can abandon EI–R entirely and choose another method of data analysis. Or, two, he can model the aggregation bias that caused test failure and in so doing avoid a second stage regression. We now discuss the latter option.

Modeling aggregation bias in the first stage of EI–R requires King's extended model (which is a one stage model, as opposed to the two stage EI–R). Recall that, in the standard implementation of King's ecological inference technique the mean of $\left(\beta_i^b, \beta_i^w\right)$ is $\left(\mathfrak{B}^b, \mathfrak{B}^w\right)$, and this mean vector is fixed across observations. However, in King's extended model the two elements of this latter vector are denoted $\left(\mathfrak{B}_i^b, \mathfrak{B}_i^w\right)$ where subscripts denote dependence on observation i.

Readers interested in the details of the extended model should consult King (1997, pp. 168–74). In brief, this model posits that $\mathfrak{B}_i^b$ and $\mathfrak{B}_i^w$ are linear functions of covariates in much the same way as EI–R posits that β_i^b and β_i^w are linear functions of covariates. That is, the extended model assumes something akin to $\mathfrak{B}_i^b = \theta_b' Z_i$ where θ_b is a vector of parameters to be estimated and, as before, Z_i is a vector of covariates. A product of the extended model, then, is an estimate $\hat{\theta}_b$ of the parameter vector θ_b (and, similarly, an estimate $\hat{\theta}_w$ of θ_w based on $\mathfrak{B}_i^w = \theta_w' Z_i$). For numerical reasons (equation 9.2, p. 170), King's extended model does not assume the precise linear form noted above for the two means $\mathfrak{B}_i^b$ and $\mathfrak{B}_i^w$. However, the expressions for these means capture the essence of the model.

The extended model allows dependence of β_i^b and β_i^w to be modeled in one stage; hence, when the extended model is employed there are no second stage regressions. If, for instance, Z_i contains only a single covariate, then θ_b (a scalar) describes how this covariate influences β_i^b and θ_w describes how it influences β_i^w. Similarly, when Z_i contains multiple covariates, then θ_b and θ_w are vectors with elements corresponding to covariates in Z_i.

When King's extended model is used, assessing the impact of covariates on β_i^b (and β_i^w, if this, too, is an object of study) requires considering first stage standard errors produced by King's technique. Namely, for each covariate used to parameterize the means $\mathfrak{B}_i^b$ of β_i^b and $\mathfrak{B}_i^w$ of β_i^w in the extended model, King's method produces two coefficient estimates (akin to elements of $\hat{\theta}_b$ and $\hat{\theta}_w$) and two associated t–statistics. Significance tests based on these t–statistics can be used to determine whether covariates are related to β_i^b and β_i^w.

4.1 Combining the Extended Model with EI–R

A researcher may be tempted to include some elements of the second stage covariate vector Z_i in both the first stage of King's extended model as well as in a second stage regression (e.g., Burden and Kimball 1998). This practice should be avoided as it can lead to internal incoherence. Specifically, as illustrated in Section 5, it is possible for a first stage extended ecological inference model to indicate that a given covariate has a significant impact on β_i^b and/or β_i^w whereas a second stage regression finds no evidence of such an effect, or vice versa. This would leave a researcher with contradictory estimates of the same quantity. Thus, to implement EI–R either King's standard model should be used followed by a second stage regression or King's extended model should be employed and there should be no second stage regression.

4.2 Dominance of the Extended Model

When our test for logical consistency of EI–R rejects when applied to a given ecological dataset, EI–R is dominated by King's extended model insofar as every problem which affects the latter also affects EI–R, plus EI–R is logically inconsistent. When one statistical procedure has associated with it a collection of problems and when a second procedure has the same collection of problems plus others, then the latter procedure should not be used.

The issue of dominance is an important one because one might think that quantitative work in political science routinely violates assumptions and that logical inconsistency in EI–R is no worse than other problems that are regularly ignored by scholars engaged in ecological inference. Forthcoming evidence in Section 5 suggests that logical inconsistency in EI–R is not a minor problem. Moreover, even if the problem were minor, researchers should always seek to make their analysis as internally coherent as possible and to violate as few assumptions as possible.

Furthermore, another reason to use King's extended model in lieu of EI–R concerns other sources of logical inconsistency that can plague EI–R. We have focused solely on whether the two stages of EI–R make contradictory assumptions about aggregation bias. But, there are other ways in which EI–R can be logically inconsistent. For instance, EI–R's first stage assumes that $\left(\beta_i^b, \beta_i^w\right)$ have a truncated, bivariate normal distribution, and this has implications for the disturbance terms ν_i^b and ν_i^w in equations (5) and (7), respectively. It is unknown whether second stage assumptions about ν_i^b and ν_i^w are compatible with first stage bivariate normality, and the use of King's extended model renders this issue moot.

Finally, and unrelated to logical inconsistency, Herron and Shotts (2001) argue that EI–R estimates are inconsistent regardless of whether the first stage of EI–R is compatible with its second. Thus, according to Herron and Shotts, logical inconsistency is irrelevant and EI–R is by construction an inaccurate procedure.

5 An Application of the Test for Logical Consistency in EI–R

We now apply our test for logically consistency of EI–R to the ecological dataset used by Burden and Kimball (1998) in their study of ticket–splitting in the United States Congressional and Presidential elections of 1992.[7] For Burden and Kimball's House election data, X_i is the proportion of House voters in Congressional district i who voted for Michael Dukakis in the 1992 presidential race, and T_i is the fraction who voted for the Democratic House candidate.[8] Furthermore, β_i^b is the fraction who voted for a Democratic Congressional candidate as well as for Dukakis (straight ticket Democrats), and β_i^w is fraction who voted for Republican presidential candidate George H. W. Bush and a Democratic Congressional candidate (Burden and Kimball call these RD ticket splitters). In their application of EI–R to the study of RD ticket splitting, Burden and Kimball use estimates $\hat{\beta}_i^w$ as dependent variables in a second stage regression. Burden and Kimball also use values of $1 - \hat{\beta}_i^b$ as dependent variables in an EI–R analysis of DR ticket splitting, i.e., rates at which individuals voted Democratic in the presidential race and Republican in a Congressional contest.[9]

The second stage covariate vector Z_i that Burden and Kimball use to study RD ticket splitting has four elements. These consist of an indicator as to whether district i had a Democratic incumbent (*Democratic Incumbent*), the Democratic candidate's proportion of House campaign spending (*Spending Ratio*), whether district i allowed straight–party voting (*Ballot Format*), and whether district i was located in the South (*South*).

To apply our specification test for logical inconsistency of EI–R to Burden and Kimball's analysis of RD ticket splitting, we regress the four elements of Burden and Kimball's Z_i vector on X_i; this produces four regressions. In terms of DR ticket splitting, Burden and Kimball

[7] Cho and Gaines (2000) identify several errors in Burden and Kimball's dataset, but we use the original dataset to maintain compatibility with the published study. Burden and Kimball's data is available from the ICPSR (ftp://ftp.icpsr.umich.edu/pub/PRA/outgoing/s1140).

[8] As Burden and Kimball note, because of rolloff X_i is not actually known and must be estimated. Like Burden and Kimball, we estimate X_i using King's ecological inference technique. As pointed out in Herron and Shotts (2001), these estimated turnout rates are contaminated by errors. However, for purposes of comparability we follow Burden and Kimball and treat these estimates as if they were true values.

[9] We comment only on Burden and Kimball's House data because their Senate election dataset is extremely small ($n = 32$). Furthermore, Burden and Kimball augment their Senate data with House data in a way that is very hard to justify.

Table 1: Testing for Logical Consistency in EI–R as Applied to Burden and Kimball (1998)

Variable	Estimate
Democratic Incumbent	1.78
	(10.6)
Spending Ratio	1.59
	(12.6)
Ballot Format	-0.110
	(0.499)
South	-1.04
	(6.87)
Republican Incumbent	-1.71
	(10.8)
n	361

Note: t–statistics in parentheses, based on heteroskedastic-consistent standard errors

use the same covariates as in their RD analysis but with a Republican incumbent indicator in place of *Democratic Incumbent*. Table 1 display results for both RD and DR specification tests (which are the same except for the incumbent issue).

Consider the top row of Table 1. The t–statistic (10.6) from a regression of "Democratic Incumbent" on X_i (percent Dukakis in District i) is significantly positive at conventional confidence levels, and this immediately implies that Burden and Kimball's EI–R analysis of RD ticket splitting is logically inconsistent. Similarly, the "Republican Incumbent" t–statistic in Table 1 is significantly negative—implying that the probability of a Congressional District's having a Republican incumbent is relatively low in Districts that have many Dukakis voters—and thus Burden and Kimball's DR ticket splitting analysis is also logically inconsistent.

In light of Burden and Kimball's Z_i vector, our results on logically inconsistency should hardly be considered surprising. For example, the regression output described in Table 1 implies that the probability of a Congressional District's having a Democratic incumbent (element of Z_i) is relatively high in Districts that have many Dukakis voters (high X_i Districts). Indeed, it would be very anomalous if, as required by Proposition 2, there were no relationship between a District's presidential vote proclivity and the partisan affiliation of its incumbent Congressional representative, i.e., it would be very anomalous if the mathematical restriction in Proposition 2 held when applied to Burden and Kimball's EI–R analysis of RD ticket splitting.

Of course, there is nothing special about incumbent indicators. In fact, practically every result in Table 1 implies that Burden and Kimball's DR and RD applications of EI–R are logically inconsistent. It is crucial to recognize that logical inconsistency follows if even one element of Z_i is mean dependent on X_i. So, when almost all specification tests reject for a given ecological dataset, i.e., when all elements of Z_i are mean dependent on X_i, then there is overwhelming evidence against EI–R.

Thus, a researcher who wants to study RD and DR ticket splitting with Burden and Kimball's dataset should use King's extended ecological inference model, a model that can be used in one stage to estimate how covariates affect ticket splitting rates. See Table 2 for results.[10] The table also reports verbatim Burden and Kimball's second stage regression estimates which, as now known, are based on a logically inconsistent application of EI–R.[11]

It is important to note that EI–R and extended model estimates for RD ticket splitting (and DR ticket splitting as well) are not directly comparable. Extended model coefficients are not regression coefficients but rather represent shifts in the mean of the distribution from which β_i^b and β_i^w are drawn; see p. 170 in King (1997) for a full explanation. However, to determine whether Burden and Kimball's overall EI–R conclusions differ from those produced by an analysis of ticket splitting that uses King's extended model, one can compare the signs and the t–statistics of the RD and DR estimates in Table 2.

In their analysis of RD ticket splitting, Burden and Kimball find that *Democratic Incumbent*, *Spending Ratio*, and *South* all have positive and significant effects on Republican

[10] Our implementation of King's extended model, which is based on standard likelihood theory, does not incorporate priors. We tried a large variety of different starting values for the extended model, found what appear to be two local maxima, and thus present results for the local maximum with the higher loglikelihood value. We are confident that this is the true function maximum since there was a huge difference between the two maxima we found and because the other local maximum was clearly substantively unreasonable; its estimated fraction of Democratic presidential voters who voted for Democratic House candidates was only slightly higher than its estimate of the fraction of Republican presidential voters who voted for Democratic House candidates.

[11] The Burden and Kimball (1998) EI–R results are from Tables 6 and 7 on pp. 539 and 540, respectively.

Table 2: King's Extended Model as Applied to Burden and Kimball's Analysis of House Elections

	RD Estimates		DR Estimates	
Variable	EI–R	Extended	EI–R	Extended
Democratic Incumbent	0.107***	0.116***	—	—
	(0.015)	(0.0273)		
Republican Incumbent	—	—	0.085***	0.151***
			(0.011)	(0.0360)
Spending Ratio	0.350***	0.0845	-0.252***	-0.588
	(0.021)	(0.324)	(0.015)	(0.375)
Ballot Format	-0.032***	0.201	0.003	0.277
	(0.008)	(0.199)	(0.006)	(0.225)
South	0.052***	0.220	0.007	0.219
	(0.009)	(0.183)	(0.006)	(0.212)

Note: standard errors in parentheses; $^{***}\,p < 0.001$, $^{**}\,p < 0.01$, $^{*}\,p < 0.05$

presidential voters' propensities to split their tickets. Moreover, Burden and Kimball find that *Ballot Format* has a negative and significant effect. In the extended model estimates for RD ticket splitting, however, only one of these results holds: *Democratic Incumbent* has a positive and significant effect on RD ticket splitting. The other three coefficient estimates, which in the logically inconsistent EI–R analysis are highly significant, are in fact statistically insignificant once this problem is corrected for.[12]

We now turn to Burden and Kimball's analysis of DR ticket splitting. In the rightmost two columns of Table 2, the logically inconsistent EI–R results indicate that *Republican Incumbent* and *Spending Ratio* have highly significant effects on DR splitting. In the extended model analysis, though, *Republican Incumbent* has a significant effect whereas there is no evidence to support the claim that *Spending Ratio* is similarly significant.

There are, notably, no results in Table 2 that are significant in an extended model analysis yet insignificant in EI–R. This means that, by employing EI–R in a logically inconsistent way, Burden and Kimball generated evidence of "findings" that, in fact, are not findings once

[12]This does not mean, of course, that Burden and Kimball's claims are unequivocally wrong. Rather, it means that there is no evidence suggesting that their claims are correct.

logical consistency is adjusted for.[13]

Thus, we conclude that Burden and Kimball's alleged substantive results on ticket splitting are not results at all and instead are artifacts of a self-contradictory statistical technique. Using a method that is logically consistent we have shown that, with one exception (the effect of incumbent partisan affiliation), *all* of Burden and Kimball's substantive findings are statistically unfounded. The exception is notable since it is quite intuitive: House districts with Democratic incumbents are relatively likely to produce split tickets that support both a Republican presidential candidate as well as the incumbent Congressional incumbent (and symmetrically for districts with Republican Congressional incumbent). Furthermore the extended model analysis of Burden and Kimball's House data returns several non-null results, and this means that its failure to do so in a number of cases does not reflect an overall inability of the extended model to estimate anything in a precise way.[14]

For instance, Burden and Kimball claim that, "The most robust finding from [various applications of EI–R] is that congressional campaign spending has a dramatic influence on the percentages of voters who split their ballot[s] (p. 542)." In fact, there is no evidence in Burden and Kimball's dataset to support this assertion. Similarly, Burden and Kimball claim that statistical significance of *South* in the RD analysis leads to a rejection of theories of intentional ticket splitting (p. 540). In fact, *South* is insignificant ($t = 1.2$) in the logically consistent RD analysis. These comments do not mean, of course, that all of Burden and Kimball's conclusions are incorrect. However, Burden and Kimball's data simply do not support their substantive claims about the causes of split ticket voting and, in particular, Burden and Kimball's dataset does not cast aspersions on theories of intentional ticket splitting insofar

[13]To confirm that the differences identified in Table 2 were due to the distinction between King's extended model and EI–R, we replicated Burden and Kimball's EI–R analysis and found the same regression results they reported in their article.

[14]Besides logically inconsistency, Burden and Kimball's RD analysis also suffers from the internal incoherence discussed in Section 4.1. This is because Burden and Kimball include *South* in their first stage ecological inference model as well as in second stage regressions. They allow, to be precise, this covariate to influence β_i^w, the probability that an individual votes RD. For reasons that are unclear, they do not allow *South* to influence $1 - \beta_i^b$, the probability that the individual votes DR. The first stage estimate of the impact of *South* on RD ticket splitting is 0.0622 with a standard error of 0.0410 ($t = 1.52$, which differs from the t-statistic in Table 2 because Burden and Kimball's first stage did not include variables other than *South*). Thus, according to Burden and Kimball's first stage there is no compelling evidence that *South* affects RD ticket splitting. But, according to their second stage, there is evidence of this.

as *South* is concerned.

Our results on Burden and Kimball's House dataset raise an important question: are problems with applications of EI–R common or is the example we use here anomalous? Of course, without an exhaustive study of numerous research projects that use second stage regressions, we will never know precisely how many published applications of EI–R are based on logically inconsistent assumptions.[15]

Nonetheless, in many cases where applied researchers use EI–R it is exceedingly hard to imagine that elements of the second stage covariate vector Z_i are related to β_i^b (as assumed in a second stage regression) yet unrelated to X_i (as assumed in first stage ecological inference). For example, in Cohen, Kousser and Sides's (1999) study of crossover voting in state legislative elections in Washington and California, X_i is the fraction of individuals in district i who are Democrats.[16] Moreover, one element of Cohen, Kousser and Sides's second stage covariate vector Z_i is an indicator for whether there is a Democratic incumbent in district i. It is in theory conceivable that Democratic incumbency in district i is unrelated to whether district i has a large number of Democrats in it. However, this seems quite unlikely, and logical consistency of Cohen, Kousser and Sides's application of EI–R must therefore be questioned.

6 Discussion

We have proposed a test for logical consistency of the statistical procedure EI–R and illustrated its usefulness by applying the test to an ecological dataset from Burden and Kimball (1998). The test easily rejects logical consistency of Burden and Kimball's EI–R analysis of the causes of ticket splitting in House elections. Moreover, in a logically consistent statistical analysis of Burden and Kimball's ticket splitting dataset, the vast majority of Burden and Kimball's findings vanish. Indeed, we suspect that many if not the vast majority of implementations of EI–R are logically inconsistent, and that the ostensible results generated by

[15] As Burden and Kimball (1998) and Gay (2001) are the most prominent, published applications of EI–R, we sought to scrutinize both of these studies for evidence of logical inconsistency. We were, however, unable to scrutinize Gay's results because she would not release her dataset to us (personal communication with Claudine Gay, 2002).

[16] Cohen, Kousser and Sides's notation uses P_i instead of X_i and V_i instead of T_i.

these implementations are artifacts of a self-contradictory statistical technique.

If our critique of EI–R is taken seriously, we should expect to see numerous researchers who wish to estimate second stage regressions instead relying on King's extended model, which does not suffer from logical consistency problems. However, Rivers (1998) notes that King's extended model is fragile in the sense of being barely identified. The importance of this point is as follows: if it is difficult to identify using King's extended model the impact of covariates on unknown, disaggregated quantities, then it should be commensurately difficult to use a second stage regression to do this. Yet second stage regressions are easily identified and they make it appear as if identifying the impact of covariates on disaggregated quantities is a simple matter. Of course, second stage regressions are identified precisely because they treat ecological inference point estimates as data.

A desire for identification might explain why researchers often turn to EI–R in the first place (of course, the identification achieved is illusory, since EI–R is identified only because it treats unknown quantities as known). Even so, one might wonder why anyone would ever estimate a second stage regression when covariates to be included in such a regression can be incorporated into a first stage extended ecological inference model.

We suspect that researchers have turned to EI–R, and may wish to continue doing so, because applications of its logical replacement, King's extended model, are likely to generate numerous null results from which few substantive conclusions about political behavior can be drawn. A prevalence of null results does not reflect flaws in the extended model. Rather, null results reflect the difficulty of incorporating covariates in ecological inference.

In contemporary political science research, statistically significant "findings" are valued more than null results. But, findings should not be taken seriously if they exist solely because of a statistical quirk, i.e., solely because they appear in a logically inconsistent application of EI–R yet not in a logically consistent ecological inference model. Some problems in political science, and the study of ticket splitting may be one of these, may simply be intractable with aggregate data, and it is better to recognize and accept this rather than attempt to extract unsupportable findings using a potentially misleading statistical technique.

References

Achen, Christopher H. and W. Phillips Shively. 1995. *Cross–Level Inference*. Chicago: University of Chicago Press.

Burden, Barry C. and David C. Kimball. 1998. "A New Approach to the Study of Ticket Splitting." *American Political Science Review* 92(3):533–44.

Burden, Barry C. and David C. Kimball. 2002. "Why Americans Split Their Tickets: Campaigns, Competition, and Divided Government." Forthcoming, *University of Michigan Press*.

Cho, Wendy K. Tam and Brain J. Gaines. 2000. "Reassessing the Study of Split–Ticket Voting." Working Paper, University of Illinois.

Cohen, Jonathan, Thad Kousser and John Sides. 1999. "Sincere Voting, Hedging, and Raiding: Testing a Formal Model of Crossover Voting in Blanket Primaries." Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.

Freedman, D.A., S.P. Klein, M. Ostland and M.R. Roberts. 1998. "On 'Solutions' to the Ecological Inference Problem." *Journal of the American Statistical Association* 93(444):1518–22.

Gay, Claudine. 2001. "The Effect of Black Congressional Representation on Political Participation." *American Political Science Review* 95(3):589–602.

Herron, Michael C. and Kenneth W. Shotts. 2001. "Using Ecological Inference Point Estimates as Dependent Variables in Second Stage Linear Regressions." Forthcoming, *Political Analysis*.

Kimball, David C. and Barry C. Burden. 1998. "Seeing the Forest and the Trees: Explaining Split–Ticket Voting within Districts and States." Paper presented at the annual meeting of the American Political Science Association, Boston, MA.

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

Lublin, David and D. Stephen Voss. 2002. "Context and Francophone Support for the Sovereignty of Quebec: An Ecological Analysis." *Canadian Journal of Political Science* 35(1):75–101.

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

Mebane, Walter R. and Jasjeet Singh Sekhon. 2002. "Coordination and Policy Moderation at Midterm." *American Political Science Review* 96(1):141–157.

Rivers, Douglas. 1998. "Review of 'A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data'." *American Political Science Review* 92(2):442.

Tam, Wendy K. 1998. "Iff the Assumption Fits...: A Comment on the King Ecological Inference Solution." *Political Analysis* 7:143–63.

Voss, D. Stephen and David Lublin. 1998. "Ecological Inference and the Comparative Method." *APSA-CP: Newsletter of the APSA Organized Section in Comparative Politics* 9(1):25–31.

Voss, D. Stephen and Penny Miller. 2000. "The Phantom Segregationist: Kentucky's 1996 Desegregation Amendment and the Limits of Direct Democracy." Working Paper, University of Kentucky.

Wolbrecht, Christina and J. Kevin Corder. 1999. "Gender and the Vote, 1920–1932." Paper presented at the annual meeting of the American Political Science Association, Atlanta, GA.