Estimating Regression Models in which the Dependent Variable Is Based on Estimates with Application to Testing Key's Racial Threat Hypothesis

Jeffrey B. Lewis
Politics Department
Princeton University *

August 4, 2000

Abstract

Researchers often use as dependent variables quantities estimated from auxiliary data sets. Estimated dependent variables (EDV) models arise, for example, in studies where counties or states are the units of analysis and the dependent variable is an estimated mean or fraction. A new source of such EDV regressions has been created by King's ecological inference estimator (King 1997). Researchers have fit regression models to quantities such as percent minority turnout that were estimated using King's EI (Gay 1998). Scholars fitting EDV models have generally recognized that variation in the sampling variance of the observations on the dependent variable will induce heteroscedasticity. In this paper, I show that the most common approach to this problem, weighting the regression by the inverses of the sampling standard errors of the dependent variable, will usually lead to inefficient estimates and underestimated standard errors. I show that the degree of this inefficiency and overconfidence can be very large. I also suggest two alternative approaches that are simple to implement and more efficient and yield consistent standard error estimates. I then apply these methods to testing the contemporary relevance of Key's racial threat hypothesis. Revisiting parish-level data from Voss (1996), I find that support for David Duke in several elections was increasing in the fraction of parish voters that are African-American supporting Key's hypothesis.

*email: jblewis@princeton.edu. I thank Michael Herron, Barry Burden, Lynn Vavreck, and Gary King for comments on an earlier drafts of this paper. I thank Steve Voss for pointing me toward the empirical example. Any remaining errors are mine alone.

It is not uncommon for regressions to be run in which observations on the dependent variable are estimated in an auxiliary analysis. For example, several researchers have recently used estimates from King's (1997) ecological inferences estimator as dependent variables in a subsequent analysis (Burden & Kimball 1998, Gay 1998, as examples).[1] Dependent variables may also be regression coefficients as, for example, in Gelman & King (1994), Martinez (1990), or Roberts (2000). Researchers often run regressions using poll question marginals such as presidential approval as dependent variables (MacKuen, Erikson & Stimson 1989, MacKuen, Erikson & Stimson 1992, Oppenheimer 1996, Taylor 1980). More generally, election vote shares, legislative seat shares, fractions of bills vetoed, and other similar potential dependent variables are best thought of as "estimated" in the sense that they are aggregated outcomes of stochastic processes (Patterson & Caldeira 1985, Voss 1996, for example). That is, vote share for a Senate incumbent in a county that cast one million votes is surely a more reliable indicator of support for the incumbent in that county than is the vote share for the incumbent in a precinct in which only 20 people voted. The fitting of such estimated dependent variable (EDV) regressions is the subject of this note.

Estimation errors in the dependent variable do not necessarily complicate regression analysis. Indeed such errors of measurement are often explicitly included in discussions of regression residuals presented in introductory textbooks (Maddala 1992, p. 65). However, if the sampling uncertainty in the dependent variable is not constant across observations, the regression errors will be heteroskedastic and OLS will not be efficient and may produce inconsistent standard error estimates.

When the dependent variable in a regression is based on estimates, the regression residual can be thought of as having two components. The first component is sampling error (the difference between the true value of the dependent variable and its estimated value). The second component is the random shock that would have obtained even if the the dependent variable was directly observed as opposed to estimated. The first component will be heteroscedastic if the sampling vari-

[1] Herron & Shotts (2000) point out that using the so-called "precinct-level" EI estimates as dependent variables in second-stage regressions will lead to attenuated estimates and more generally calls into the question the validity of using "precinct-level" EI estimates in subsequent analysis. As stated in the text, this paper is addressed to a far more general problem of which the use of EI estimates as dependent variables is but one example. However, it should be noted that the techniques that I present below are predicated on the assumption that the data used are free of the features described by Heron & Schotz. In particular, I assume that the sampling or measurement error in the dependent variable ($Y^* - Y$) is independent of the independent variables and error term ($\mathbf{X}$ and ϵ) of the regression.

ance differs across observations.[2] However, the second component could well be homoscedastic. The usual approaches to the problem are either to run OLS ignoring heteroscedasticity resulting from the first component of the residual or to run WLS assuming the entire residual, and not just the first component, is heteroscedastic. Both of these approaches are inefficient and to produce inconsistent estimates of parameter uncertainty.

In particular, I demonstrate the following intuitive relationships in monte carlo experiments:

1. The larger the share of the regression residual that is due to sampling error in the dependent variable, the less efficient OLS is and the more efficient WLS is. However, and perhaps more surprisingly, only when the share of the regression residual due to sampling error in the dependent variable is very high (roughly 90 percent or more) does WLS produce more efficient estimates than OLS.

2. WLS leads to greater overconfidence (downwardly biased standard errors) the smaller is the fraction of the error variance attributable to sampling error in the dependent variable.

3. OLS standard errors are increasingly biased as the correlation between the independent variables and the variance of the sampling error in the dependent variable increases and as the fraction of the total regression error variance due to sampling error in the dependent variable increases.

4. White's (1980) heteroscedastic consistent standard error estimator yields generally good results (no over or under-confidence), though as mentioned above OLS may be quite inefficient.

5. The simple two-stage estimators described below produce efficient estimates (relative to OLS) if a sufficiently high fraction of the total regression error variance is due to sampling error. The standard errors of these estimators generally produce less overconfidence than is found using OLS and WLS.

[2] For example, reported average state income may be based on a much larger sample in California than it is in Rhode Island leading California's mean income to be more precisely estimated than Rhode Island's. Some surveys such as the 1988 Senate Election Study intentially draw samples of roughly equal size from each aggregate unit, thus avoiding much of the heteroscedasticity that is generally present when sample means are used as a dependent variable. In these cases, the hetroscedasticiy from sampling error would only enter from inter-unit heterogeneity in the intra-unit variance of the variable being sampled.

Given the drawbacks of the conventional approaches to this problem, I present two alternative FGLS estimators. Both of these approaches allow the analyst to address the problem of heteroskedasticity in the first error component without assuming that the second component is similarly heteroscedastic. The first approach assumes that the sampling variances of the observations on the dependent variable are known as would be the case if, for example, the data were regression estimates or vote aggregated survey responses. The second estimator requires only that the sampling errors be known up to a proportional factor as would be the case if the data were means based on samples of differing sizes.[3] Both of these estimators are very easy to implement using standard statistical packages.[4] I will show in monte carlo experiments that both of these estimators are more efficient than OLS and the traditional WLS approaches in most cases.

The paper is organized as follows. Section 2 formalizes the problem of regressions involving estimated dependent variables. Section 3 describes the monte carlo experiments. Section 4 presents an empirical example. In the example based on Voss (1996), a regression model of parish-level support for David Duke in three Louisiana elections supports V.O. Key's (1949) racial threat hypothesis if estimated by OLS, but not if estimated by the standard WLS approach. Using the more appropriate estimators presented in section 2, Key's hypothesis is supported. Section 5 concludes.

1 Regressions with estimated dependent variables

Consider the following regression model,

$$y_i = \beta_1 + \sum_{j=2}^{k} \beta_j x_{ij} + \epsilon_i \tag{1}$$

for observations $i = 1, \ldots, N$. Assume that this regression model meets all of the usual Gauss-Markov assumptions. However, y_i is not observable. Rather we observe an unbiased estimate y_i^* where

$$y_i^* = y_i + u_i \tag{2}$$

[3] In this case, if the within unit variance of the variable were constant, the sampling variances would be proportional to $1/n_i$ where n_i is the size of the sample from which the mean was calculated for observation i.

[4] A STATA command implementing both of these procedures is available from the author.

and $\mathrm{E}(u_i) = 0$ and $\mathrm{Var}(u_i) = \omega_i^2$ (for $i = 1, \ldots, N$). u_i is the sampling error in y_i^* and ω_i^2 is the variance of that sampling error. Assume that the estimate of the dependent variable for each observation is independent of the others.[5] In particular, assume $\mathrm{Cov}(u_i, u_j) = 0$ for all $i \neq j$. Now consider the regression formed by substituting equations 2 into equation 1,

$$y_i^* = \beta_1 + \sum_{j=2}^{k} \beta_j x_{ij} + u_i + \epsilon_i.$$

Writing $v_i = \epsilon_i + u_i$, we have

$$y_i^* = \beta_1 + \sum_{j=2}^{k} \beta_j x_{ij} + v_i. \tag{3}$$

Obviously, if $\omega_i \neq \omega_j$ for some i and j, v_i is heteroskedastic. In particular, letting $\mathbf{v} = (v_1, v_2, \ldots, v_N)'$,

$$\mathrm{E}(\mathbf{v}\mathbf{v}') = \mathbf{\Omega} = \begin{bmatrix} \sigma^2 + \omega_1^2 & 0 & \ldots & 0 \\ 0 & \sigma^2 + \omega_2^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \sigma^2 + \omega_N^2 \end{bmatrix} \tag{4}$$

where $\mathrm{Var}(\epsilon) = \sigma^2$. If σ^2 and ω_i^2 were both known, a WLS approach would yield best linear estimators. That is, OLS estimation of

$$w_i y_i = \beta_1 w_i + \sum_{j=2}^{k} \beta_j x_{ij} w_i + w_i v_i$$

where

$$w_i = \frac{1}{\sqrt{\sigma^2 + \omega_i^2}} \tag{5}$$

yields efficient estimates of the regression parameters.

In real-world applications, we generally only have estimates of the ω_i^2 ($i = 1, \ldots, N$) and no knowledge of σ^2. In the next section, I consider four feasible approaches to this problem in cases in which knowledge or σ^2 and/or the ω^2's is missing or incomplete.

[5]The first FGLS approach described below is trivially extended to the case in which the estimates are not independent.

2 Approaches to fitting estimated dependent variable regressions

I will consider four approaches to estimating EDV regressions. The first two are the most commonly applied: OLS (with or without White's heteroscedastic consistent standard errors) and WLS using weights based on the inverse of the standard errors of the dependent variable estimates. The third is a two-stage procedure first suggested in Hanushek (1974) appropriate for cases in which estimates or the true sampling variances of the y_i^* are available (i.e., the ω^2's are known). The forth is also a two-stage procedure appropriate for cases in which the sampling variance of the y_i^*'s are known only up to a constant of proportionality.

2.1 Estimation by Ordinary Least Squares

Given the assumptions made previously, it is well-known that equation 3 can be consistently estimated by OLS. However, in order for OLS to be efficient, ω_i must be constant across all observations. That is, only if the sampling variances of the y^* are constant (or zero) will v_i be homoscedastic and OLS efficient.

In general, OLS will not be efficient and the usual standard error estimator is, in general, inconsistent. Thus, OLS estimates will be less precise than is optimal and, under some conditions, will produce badly inconsistent standard error estimates. Not surprisingly, if the degree of variation in ω_i^2 is small or if σ^2 is large relative to the ω^2's then OLS will perform quite well. That is, if the appropriate set of weights (w_i for $i = 1, \ldots, N$) are nearly constant across observations, OLS will be nearly efficient.

The inconsistent OLS standard errors can be corrected using White's heteroscedastic consistent standard error estimator(White 1980). As will be shown below, White's robust standard error estimator will correctly measure the uncertainty in OLS even in fairly small samples. However, the OLS estimator may be quite inefficient in some cases. OLS with White's standard errors is inefficient because the partial information about the nature of the heteroscedasticity (that is, knowledge of ω_i^2's) is not used.

2.2 *Estimation by Weighted Least Squares*

The most common WLS approach to EDV models is simply to set $w_i = 1/\omega_i$ ($i = 1, \ldots, N$). Such weighting is recommended in Saxonhouse (1976) and King (1997, p. 290) among other places. As shown in equation 5, this weighting scheme implicitly sets $\sigma^2 = 0$. In other words, when weighting by $1/\omega$, researchers are assuming that the R^2 in their regression would be 1 ($\sigma^2 = 0$) if they directly observed the true y's as opposed to the estimated y^*'s.

Of course, the assumption that the regression R^2 would be 1 if the true dependent variable was directly observable seems somewhat farfetched. However, as will be shown in the monte carlo simulation in the next section, if σ^2 is very small in comparison to the ω's, the assumption that $\sigma^2 = 0$ may not be too costly. However, as σ^2 grows relative to the ω_i's, the WLS estimator is badly inefficient and generates highly misleading standard error estimates.[6]

Given only these two options, researchers face a trade-off. Use OLS with robust standard error estimates and get possibly inefficient estimates, but good estimates of the parameter uncertainty. Or, use WLS and get possibly much more efficient estimates, but risk getting highly inefficient estimates and very misleading standard errors. Some scholars have, in fact, presented both estimates (Burden & Kimball, 1998). This is a good practice. However, when the methods yield estimates that differ in substantively significant ways simply reporting both sets of results is not entirely satisfying.

In the next two subsections, I will present two alternative estimation techniques that are asymptotically efficient and whose standard errors are consistent regardless of the relative size of σ^2 and the ω^2's. Both of these estimators are FGLS estimators that use OLS to generate consistent estimates of the v's. These estimated v's are used to find consistent estimates of $\omega_i^2 + \sigma^2$ (for $i = 1, \ldots, N$) from which weights are created and a second stage WLS model fit.

2.3 *Estimating $\omega^2 + \sigma^2$ when ω^2 is known*

In this section, I present a FGLS estimator that is appropriate for cases where the standard errors of the estimated dependent variable are known, or at least where reliable estimates are available.

[6] These conclusion are not specific only to EDV case, but generalize to other cases in which WLS is applied using incorrect weights (see Greene (1999, p. 505)).

This method was first suggested by Hanushek (1974) for the case where the dependent variable is estimated regression coefficients. Since the ω's are assumed to be known, only an estimate of σ^2 (the variance of the component of the regression residual that is not due to sampling of the dependent variable) is required in order to construct a proper set of weights to use in a second-stage WLS regression. To arrive at an estimate of σ^2, first run OLS and calculate residuals $\hat{v}_i$ for $i = 1, \ldots, N$. Following Hanushek (1974) the expectation of the sum of squared residuals from this OLS regression can be written as,

$$\mathrm{E}\left(\sum_i \hat{v}_i^2\right) = \mathrm{E}(\mathbf{v}'\mathbf{v}) - \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}\mathbf{X}\right)$$

where $\mathbf{\Omega}$ is the variance covariance matrix of the vector of regression residuals $\boldsymbol{v}$ given in equation (4) and $\mathrm{tr}(\cdot)$ is the trace operator.[7] $\mathrm{E}(\mathbf{v}'\mathbf{v}) = \mathrm{E}(\sum_i v_i^2)$ and

$$\begin{aligned}\mathrm{E}(\textstyle\sum_i v_i^2) &= \mathrm{E}[\textstyle\sum_i(\epsilon_i^2 + u_i + 2\epsilon_i u_i)] \\ &= \mathrm{E}[\textstyle\sum_i \epsilon_i^2] + \mathrm{E}[\textstyle\sum_i u_i^2] + 0 \\ &= N\sigma^2 + \textstyle\sum_i \omega_i^2.\end{aligned}$$

Writing $\mathbf{\Omega}$ as $\sigma^2\mathbf{I} + \mathbf{G}$ where $\mathbf{I}$ is and $n \times n$ identity matrix and $\mathbf{G}$ is an $n \times n$ diagonal matrix with ω_i^2 as the ith diagonal element, we have

$$\begin{aligned}\mathrm{E}\left(\textstyle\sum_i \hat{v}_i^2\right) &= N\sigma^2 + \textstyle\sum_i \omega_i^2 - \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\sigma^2\mathbf{I} + \mathbf{G})\mathbf{X}\right) \\ &= N\sigma^2 + \textstyle\sum_i \omega_i^2 - \sigma^2\mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\right) - \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}\mathbf{X}\right) \\ &= (N-k)\sigma^2 + \textstyle\sum_i \omega_i^2 - \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}\mathbf{X}\right)\end{aligned}$$

where k is the number of columns in $\mathbf{X}$. Rearranging, we have

$$\sigma^2 = \frac{\mathrm{E}\left(\sum_i e_i^2\right) - \sum_i \omega_i^2 + \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}\mathbf{X}\right)}{N-k}$$

implying that

$$\hat{\sigma}^2 = \frac{\sum_i e_i^2 - \sum_i \omega_i^2 + \mathrm{tr}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}\mathbf{X}\right)}{N-k}$$

[7] The trace of square matrix is the sum of its diagonal elements.

is an unbiased estimator for σ^2. In small samples, σ^2 may be estimated to be less than 0. In these cases, σ^2 can be set equal to 0.

Given this estimator for σ^2, a set of weights w_i (for $i = 1, \ldots, N$) can be constructed such that,

$$w_i = \frac{1}{\sqrt{\omega_i^2 + \hat{\sigma}^2}}.$$

The main regression is then refit using these weights. Estimates from this regression will be asymptotically efficient.

2.4 *Estimating $\omega^2 + \sigma^2$ when ω^2 is known to constant multiplicative factor*

Sometimes researchers encounter situations in which they may only know the ω^2's up to some unknown constant of proportionality. A common example here would be a dependent variable based on sample means. That is, $y_i^* = \sum_{m=1}^{n_i} Y_{im}/n_i$. Often, in such cases, we know the size of the samples upon which the estimates are based, but we do not know the exact sampling variances of the estimated means. If we assume that the variance of Y_{im} is constant for all i and j, then we have

$$\omega_i^2 = \frac{\kappa}{n_i}$$

where κ is the variance of Y and n_i is the size of the sample upon which the ith sample mean is based.[8]

[8] The usual WLS approach described above is often advocated for this case (Hanushek & Jackson 1977). The justification is as follows. Suppose that all the variables are measured as sample means. Then assume there is an underlying individual-level regression model,

$$Y_{im} = \beta_1 + \sum_{j=2}^{k} \beta_k x_{imj} + \epsilon_{im}.$$

Assuming that ϵ_{im} is i.i.d. and taking means within each sampling unit i, we find,

$$y_i^* = \beta_1 + \sum_{j=2}^{k} \beta_k \bar{x}_{ij} + \bar{\epsilon}_i.$$

where the variance of $\bar{\epsilon}_i$ will be proportional to n_i (the number of individuals sampled in each unit i). In this case, weighting the entire regression residual by $\sqrt{n_i}$ would be appropriate, However, if some x's which are uncorrelated to ϵ at both the individual and aggregate level are omitted, we will have an additional error component that is not heteroscedastic in proportion to $1/n_i$.

Again, we first run OLS and use the estimated residuals to form a set of weights, w_i ($i = 1, \ldots, N$) that can then be used in a second-stage WLS regression. We begin by noting that, the expectation of the true squared error terms is

$$\mathrm{E}(v_i^2) = \sigma^2 + \frac{\kappa}{n_i}.$$

If $\mathbf{v} = (v_1, \ldots, v_N)'$ was observed directly, σ^2 and κ could be estimated efficiently by the OLS regression of v^2 on $(1/n)$. Since, the $\hat{v}$'s are consistent estimators for the v's, we have

$$\operatorname*{plim}_{N\to\infty} \mathrm{E}(\hat{v}_i^2) = \sigma^2 + \kappa \left(\frac{1}{n_i}\right).$$

One can then consistently estimate σ^2 and κ by regressing $\hat{v}$ on $(1/n)$. As above, this method can yield estimates of $\sigma^2 < 0$. In these cases, the regression of $\hat{v}^2$ on $(1/n)$ should be rerun constraining the constant ($\hat{\sigma}^2$) to be 0. Predicted values from the regression of $\hat{v}^2$ on $(1/n)$ are estimates of $\sigma^2 + \omega_i^2$. Using these estimates, we next construct a set of weights where

$$w_i = \frac{1}{\sqrt{\hat{\sigma}^2 + \hat{\kappa}(1/n_i)}}.$$

The general procedure is a follows:

1. Regress $\mathbf{y}^*$ on $\mathbf{X}$ by OLS and calculate the squared residuals $\hat{v}_i^2$ for $i = 1, \ldots, N$.

2. Regress $\hat{v}^2$ on $\tilde{\omega}_i^2$ where $\tilde{\omega}_i^2$ is proportional to the variance of u (i.e. $\tilde{\omega}_i^2 = \omega_i^2/\kappa$). If the constant in this regression in negative ($\hat{\sigma}^2 < 0$), the regression is rerun constraining the constant to be 0. Calculate predicted values from this regression. These predicted values are consistent estimates of $\omega_i^2 + \sigma^2$.

3. Fit the WLS regression of $\mathbf{y}^*$ on $\mathbf{X}$ using weights, $w_i = \frac{1}{\sqrt{\hat{\sigma}^2 + \hat{\kappa}\tilde{\omega}_i^2}}$.

Because this method involves estimating not only σ^2. but also κ, it will yield less efficient estimates than those that can be achieved if the ω^2's are known exactly and only σ^2 must be estimated.

Parameters for the Monte Carlo experiments

Parameter	Description	Values
C	Share of variation in the regression error due to estimation error in the dependent variable.	(0.1, 0.2,..., 1.0)
θ	Degree of variation in dependent variable sampling variances across observations.	(0.2, 0.5, 0.8)
ρ	Correlation between X and sampling variances of the dependent variable.	(0.0, 0.5, 1.0)
N	Number of observations.	500
β	Regression slope	1

Table 1: *Parameters that were manipulated in the monte carlo experiments and the values to which they were set. For each combination of parameter values, 1000 simulations were performed.*

3 Monte Carlo Analysis

In this section, I describe a series of monte carlo experiments that I conducted in order to ascertain the small sample properties of the each of the estimators for the EDV model described above.

The regression model used for all the experiments is

$$y_i = 1 + 1x_i + v_i$$

where $v_i = \epsilon_i + u_i$. Both ϵ and u are drawn from independent (conditional) normal distributions, $\epsilon_i \sim N(0, \omega_i^2)$ and $u_i \sim N(0, 1 - C)$ for $i = 1, \ldots, N$. The ω^2's are drawn from a gamma distribution. In particular, $\omega_i^2 \sim Gamma(C/\theta, 1/\theta)$.[9]

Letting u be distributed as a mixture of normals where the normal distribution's variance parameter is distributed Gamma is convenient because, as shown by Patel & Read (1996, p. 31), the variance of w_i^2 will be,

$$\text{Var}(\omega_i^2) = C\theta.$$

[9]The parameterization of the Gamma distribution used here follows DeGroot (1994). I define the density of gamma distribution as $f(z|\alpha, \beta) = [\Gamma(\alpha)]^{-1}\beta^\alpha z^{\alpha-1}e^{-\beta z}$ for $\alpha > 0$, $\beta > 0$, and $z > 0$. Given this parameterization, $\text{E}(Z) = \alpha/\beta$ and $\text{Var}(Z) = \alpha/\beta^2$. In the simulations, the density of ω_i^2 is $f(\omega_i^2|C/\theta, 1/\theta)$.

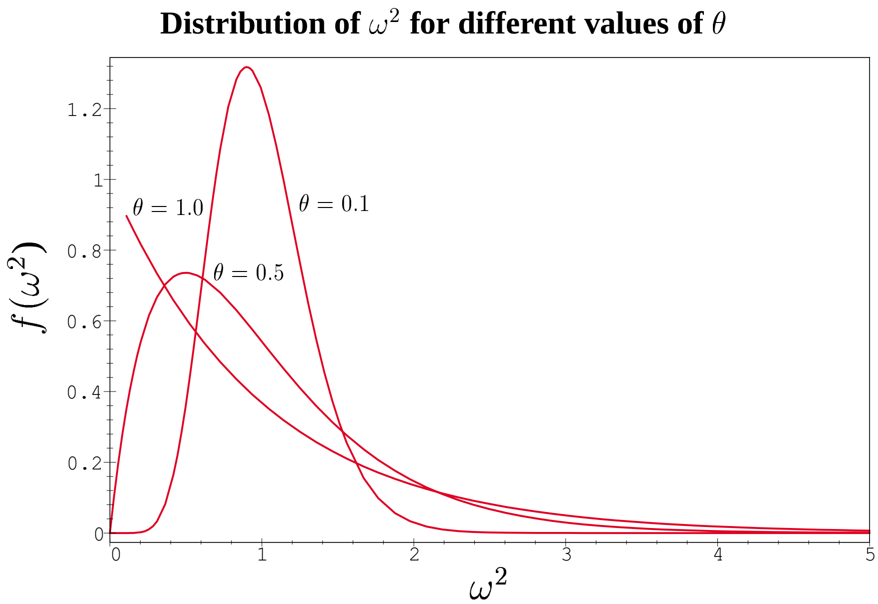

Figure 1: *Shows the distributions of ω^2s (sampling variances of the observations on the dependent variable) in the monte carlo experiments. Notice that as θ gets smaller, the frequency of extreme outliers decreases.*

Therefore, the total variance of u_i (without conditioning ω_i^2) is

$$\text{Var}(u_i) = \frac{C/\theta}{1/\theta} = C.$$

Thus, across the entire sample, $\text{Var}(v) = 1$ and C represents the share of the total regression error that is due to sampling error in the dependent variable.[10] The variance of X is also set to 1, so the R^2 in the regressions is approximately 0.5 in all cases.[11] The parameter θ describes the degree of dispersion in the ω^2's across observations. As shown in figure 1, the larger is θ, the more dispersed are the ω^2's. That is, the larger is θ, the more variation there is in the accuracy with which the dependent variable is estimated across observations. As θ approaches 0, the distribution of ω_i collapses to a point. As noted above, if ω_i is constant across observations, the error term in the regression will not be heteroscedastic and OLS regression will be efficient.

The parameters, the features of the data they relate to, and values they are set to in the experi-

[10]Note that $v_i = u_i + \epsilon_i$. Because ϵ_i is assumed to be independent of u_i, $\text{Var}(v_i) = \text{Var}(u_i) + \text{Var}(\epsilon_i) = C + (1 - C) = 1$.

[11]The "explained" variance will be $\beta^2\text{Var}(X) = 1$ and the "unexplained" variance $\text{Var}(v) = 1$, thus the R^2 will be approximately $1/(1+1) = 1/2$.

ments are shown in table 1. Before describing the results, I will first review our expectations about how each of the parameters will affect the four estimators for the EDV model developed in section 2.

When C is close to zero, sampling error in the dependent variable makes only a small contribution to the overall residual. In this case, the overall residual will be nearly homoscedastic and OLS should perform quite well. On the other hand, a WLS regression that weights by $1/\omega$ will "overcorrect" the heteroscedasticity that is generated by estimation errors in the dependent variable. Thus, *ceteris paribus*, low values of C should lead to poor WLS estimates.

The larger is θ the more heteroscedasticity is introduced by the sampling errors in the dependent variable. When θ is small, we expect that OLS will produce good estimates, because the regression error will exhibit little heteroscedasticity even if C is large. However, if θ is large, OLS is expected to be inefficient particularly as C increases.

The value of ρ is mainly expected to affect the accuracy of the standard errors. It is well known that OLS will only produce inefficient estimates and inconsistent standard error estimates if the heteroscedasticity of the regression error depends on the independent variables in some way. Thus we expect that OLS will be relatively efficient if ρ is small. However, as ρ grows, the efficiency of the OLS estimates should decline and the standard error estimates worsen.

OLS combined with White's (1980) standard errors should produce reasonable estimates of model uncertainty for all parameter settings though the estimates may be inefficient. The FGLS estimators should be efficient and produce good standard error estimates for all parameter values.[12]

In all of the experiments described below use data sets with 500 observations. Similar results are obtained in experiments using data sets with 100 and with 2500 observations. With 100 observation data sets the results were a bit noisier and, with 2500 observation data sets they are a bit cleaner. For each combinations of the parameter values 1000 simulations were preformed.

Table 2 shows the results for one particular set of parameters. In this experiment, $C = 0.7$ meaning that 70 percent of the variance in the residual is due to sampling error in the dependent variable. The variation in the sampling error of the dependent variable across observations is high ($\theta = 0.8$). In particular, across the 1000 simulations the largest sampling variance was about 40

[12]Exceptions to this claim are $C = 0$ and $C = 1$ where OLS and WLS respectively would be efficient.

Monte Carlo results for various EDV estimators

	Regression slope ($\hat{\beta}$)			
Method	Mean	Observsed mean square error (MSE)	Mean estimated standard error	Mean estimated SE as a percentage of MSE
OLS				
Usual SEs	1.000	0.051	0.045	88.2
White's SEs	1.000	0.051	0.050	98.0
WLS	0.996	0.112	0.034	30.3
FGLS				
Known variance	0.999	0.041	0.041	100.0
Proportional variance	1.000	0.043	0.041	95.3

Table 2: *Shows the results of one monte carlo experiment described in the text. The numbers shown in the table are based on 1000 simulations. The parameters values are:* $\beta = 1$, $\rho = 0.5$, $\theta = 0.8$. *and* $C = 0.7$. *Even when 70 percent of the regression residual is due to estimation error in the dependent variable, the usual WLS approach is much less efficient than OLS or the alternative FGLS approaches. The observed standard error of* $\hat{\beta}$ *is more than twice as large using WLS as it is using OLS. Moreover, WLS and, to a lesser extent, OLS systematically underestimate the uncertainty in* $\hat{\beta}$.

times as large as the smallest on average. The correlation between the sampling variances and the independent variable is set to 0.5.

All of the estimators yield estimates that are very accurate on average. Across the 1000 simulated data sets, the average estimated regression slope was nearly identical (to four decimal places) to the true value of β. This is not surprising given that all the estimators tested are known to be consistent, but it is comforting that they show little to no bias in samples of 500. This is generally true for all of the combinations of parameter values given in table 1. Indeed, across all of the experiments that I conducted the mean $\hat{\beta}$ did not differ from the true β in a statistically significant way. One convenient consequence of all of the estimators showing little to no bias it that their observed Mean Square Errors (MSEs) are basically equivalent to their observed sampling standard deviations. Thus the observed MSEs can be compared to the mean estimated standard errors to assess the degree to which each model accurately represents the uncertainty of its estimates.

While the parameter values used to generate the results in table 2 would seem to approximate the conditions under which the WLS approach would be effective, the experimental results suggest otherwise. Even with a very highly dispersed ω^2's and most the residual resulting from sampling error in the dependent variable, OLS produced more efficient estimates with more accurate standard errors than did the standard WLS. The FGLS estimators are both about 15 to 20 percent more efficient than OLS.[13] Additionally, both of these estimators produced accurate standard error estimators in this experiment.

WLS produced highly misleading standard error estimates. While the standard deviation of the WLS slope estimator across the 1000 repetitions was 0.112 (see the "MSE" column in the table), the average standard error estimate across the 1000 repetitions was 0.034. Thus, on average, the estimated standard errors were only 1/3 as large as the observed standard deviation of the estimated β across the 1000 simulations.[14] As we expect, OLS also produces inconsistent standard error estimates. However, the mean estimated OLS standard error was only 12 percent smaller than the observed OLS mean square error. However, accurate estimates of OLS parameter uncertainty were obtained using White's (1980) heteroscedastic consistent standard error estimator. As expected,

[13] That is compared to OLS, the observed standard errors are about 15.7 percent smaller for the proportional variance estimator and 19.6 percent smaller for the known variance estimator.

[14] This measure is the reciprocal of the "over-confidence" measure used by (among others) Beck & Katz (1995).

both of the FGLS estimators produced good estimates of parameter uncertainty.

Overall, in this experiment, WLS was clearly inferior to OLS and the FGLS estimators. Even with 70 percent of the total regression error resulting sampling error in the dependent variable and vast dispersion in the sampling variances across observations, OLS is preferable to the standard WLS approach. OLS with White's standard errors produced quite satisfactory results. However, the FGLS estimators, which use the information about the sampling errors in the dependent variable, did produce 15 to 20 percent efficiency gains over OLS. Whether these same conclusions hold more generally is the question to which I now turn.

Figure 3 graphs the observed standard deviations of the estimates (MSEs) as a function of the percent of the total regression error that is due to sampling the dependent variable for various values of θ and ρ. The results are consistent with expectations. When θ is low—when variance of the sampling errors in the dependent variable fairly constant across observations—all of the methods perform very similarly. Indeed, the lines representing each of the methods in the top three panels of the graph are difficult to distinguish. As the variation in the dependent variable sampling errors across observations (θ) increases, differences in the behavior of the estimators becomes apparent. In particular, when θ is large and little of the total regression residual stems from the mis-measurement in the dependent variable (C is small), WLS produces estimates with as much as 5 times the MSE as OLS. Even as this fraction is increased, OLS continues to outperform WLS. Indeed until about 90 percent of the total error variance is the result of sampling error in the dependent variable, OLS produces more efficient estimates than does the standard WLS approach.

As anticipated, the FGLS estimators produce efficient estimates relative to OLS and WLS though in many cases the gains are quite modest. The "known variance" FGLS estimator—the estimator that requires the sampling variances of the dependent variable observations be known—is generally more efficient than the proportional variance estimator. This result follows from the fact that proportional variance estimator requires the estimation of one more parameter than does the known variance estimator.

Figure 3 graphs the average estimated standard errors of the various estimators as a fraction of the observed standard errors across the 1000 experiments against the percent of the total regression error that is due to sampling error in the dependent variable. Overall, these graphs are consistent with the expectations laid out above. OLS tends to produce biased standard error estimates for high

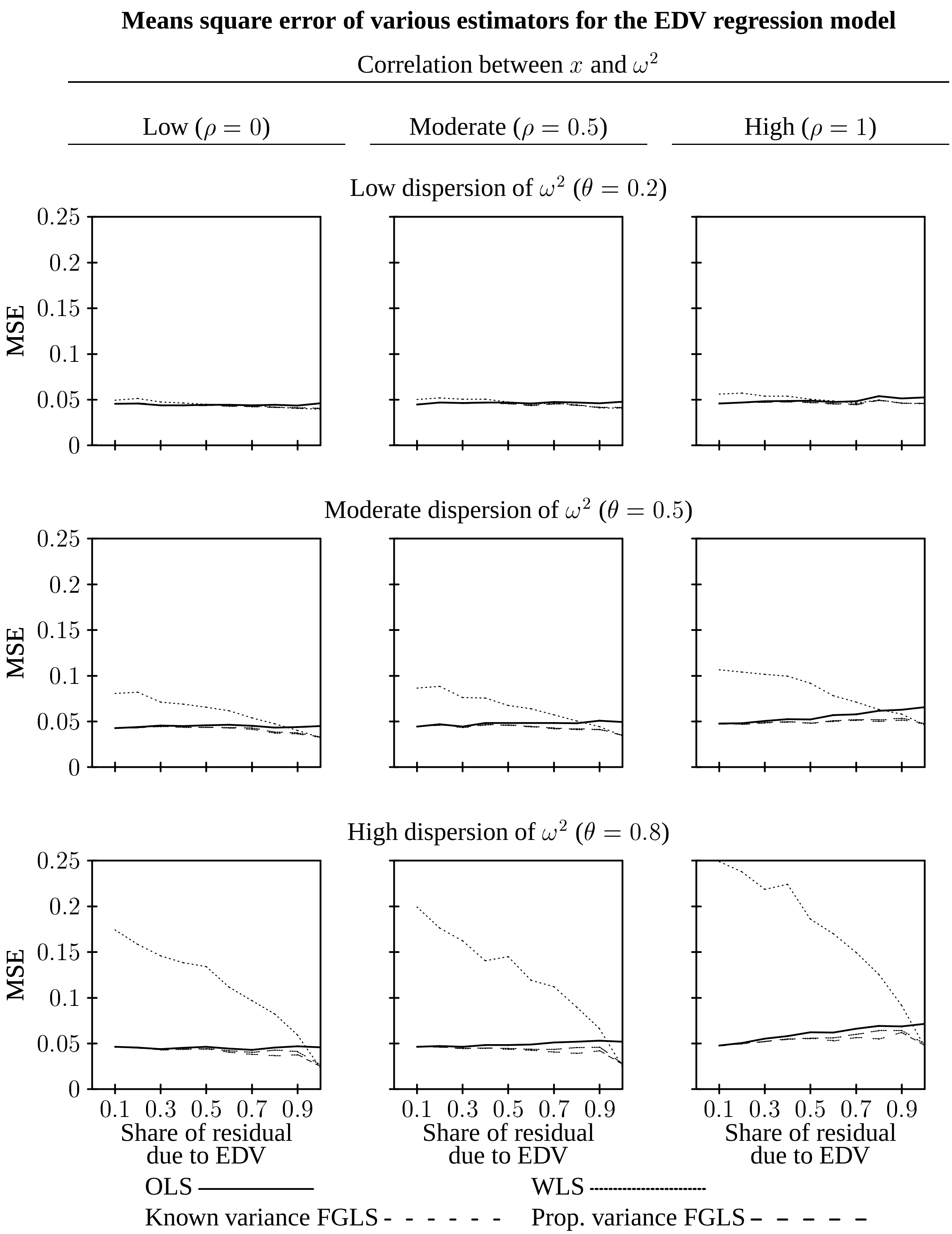

Figure 2: *Shows the mean square error of the estimated regression slope coefficient. Notice that WLS can be particularly inefficient when very little of the regression residual is due to estimation error in the dependent variable.*

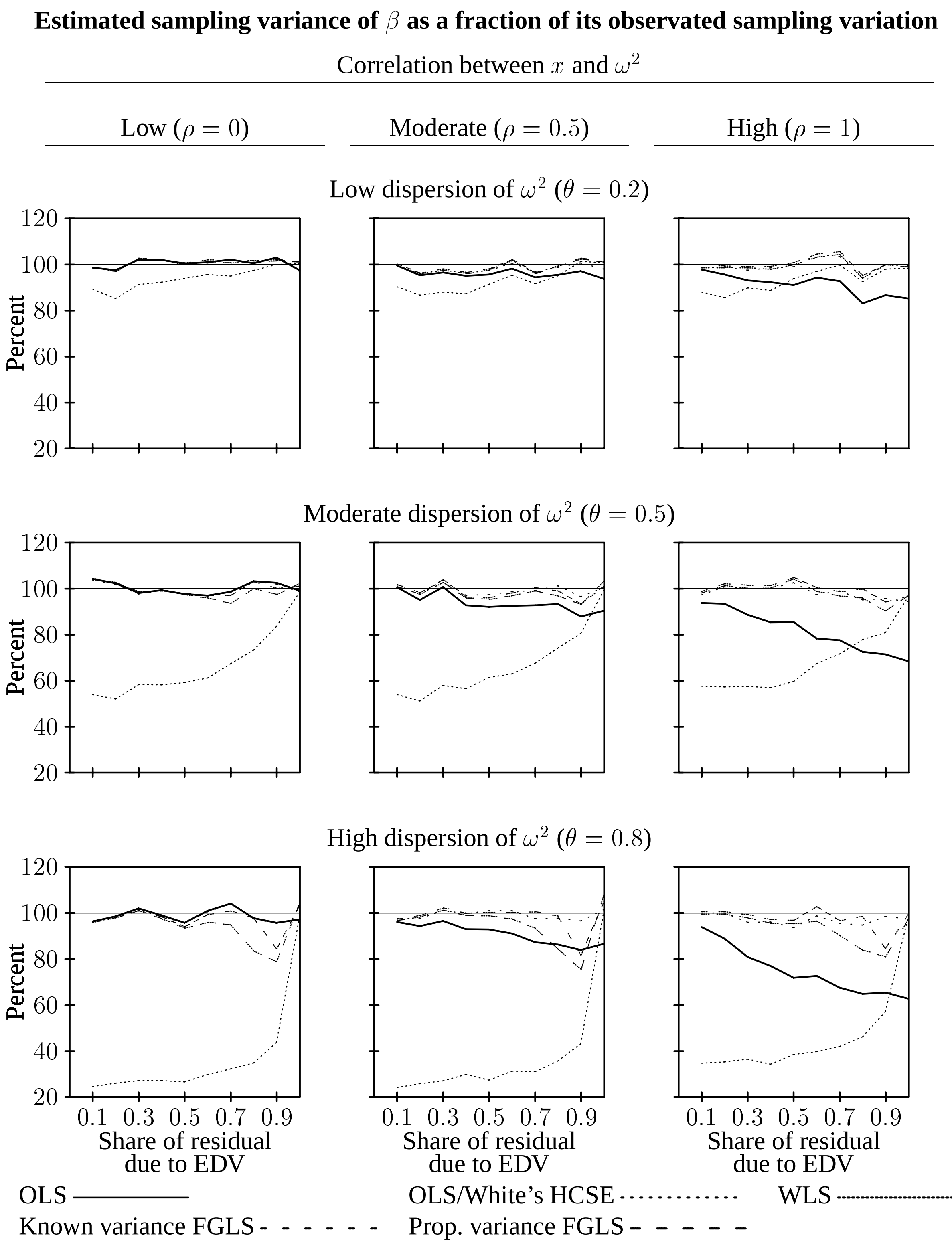

Figure 3: *Shows the estimated sampling variance of the regression slope as a percentage of its observed sampling variance. Notice that under some conditions OLS and WLS standard errors can be very misleading.*

values of C and WLS produces biased standard error estimates except when C is very high (over 0.9). White's standard error estimators is generally effective. The two FGLS estimators produce quite accurate estimates of parameter uncertainty with the exception of a few cases in the lower right corner of figure 3. High values of ρ and θ and values of C between 0.7 and 0.9, the FGLS standard error estimates are drop to as little as 80 percent of the true uncertainty.[15]

Moving from the left-most to right-most panels of figure 3 one sees that as the correlation between the independent variable and the sampling variances of the dependent variable increases the bias of the OLS standard errors increases.[16] However, White's robust standard errors are not similarly effected. It is interesting to note that even in the low sampling error dispersion case ($\theta = 0.2$), the estimated OLS and WLS standard errors are in some cases as little as 80 percent of the observed standard errors. While all the methods are quite similar in terms of their efficiency when $\theta = 0.2$ (see figure 3), the FGLS estimators or OLS with White's standard errors are still preferable because they result in better estimates of parameter uncertainty.

4 Emperical Application: Testing V.O. Key's Racial Threat Hypothesis

Several years ago there was a controversy in the in *Journal of Politics* about the relevance of Key's (1949) racial threat hypothesis to contemporary Southern politics. Key's hypothesis asserts that support for hard-line anti-black politicians is strongest among those white voters that live in areas with high concentrations of African Amercian voters (Key 1949). The logic of this argument is the following,

> In the grand outlines of politics in the South revolves around the position of the Negro. . . The core of the political South . . . is made up of those counties and sections of the southern states in which Negroes constitute a substantial proportion of the population. In these [black belt] areas a real problem of politics. . . is the maintenance of control by a white minority (Key 1949, p. 5 cited in Giles & Buckner 1993, p. 702).

By many accounts and as measured by several indicators, the politics of the South has changed considerably since 1949. Given these changes, Giles & Buckner (1993) asked whether racial threat was still an important and relevant political consideration in the post Voting Rights Act South.

[15] This occurs because for these parameter values with samples of 500, σ^2 is often estimated (by constraint) to be 0 in which case the FGLS and WLS estimator are equivalent.

[16] This result is similar to those typically found in the heteroscedasticity literature (Greene 1999, pp. 505).

To this end, they modelled parish-level support for David Duke among white voters in the 1990 Louisiana Senate primary as a function of the percentage of the all parish voters that were black.[17] They found that "[d]espite controls for urbanism, social status, levels of unemployment, in-migration, and age, support for Duke was... linked positively to the level of racial concentration in the local context" and concluded that their "analysis supports the continued potential for the operation of racial threat in the 'transformed' South" (Giles & Buckener, p. 710–11).

Revisiting Duke support in the 1990 Senate primary and expanding the analysis to include Duke support in the 1991 Gubernatorial primary and general elections, Voss (1996) raises several methodological concerns about Giles & Buckner's original analysis. Most importantly for my purposes, Voss criticized Giles & Buckner for estimating their regression models by OLS. Voss argued that OLS was inappropriate because "unlike voting districts, parishes have... highly varied numbers of registered voters," (Voss 1996, p. 1164). He concluded that "[t]he variance of each observation is not constant, but probably is inversely proportional to the number of units [individuals] within that group [parish], [thus] we can use Generalized Least Squares [more precisely, Weighted Least Squares]" (Voss 1996, p. 1166). What Voss seems to have had in mind was that each parish had some underlying level of support for David Duke that can be estimated by the observed vote share. In essence, parish vote share was treated as a large sample. Voss goes on to fit his models by WLS using $\sqrt{1/n}$ as the weight where n is the number white registered voters in a given parish. Using WLS, Voss found little support for the racial threat hypothesis concluding that his "findings cast doubt on the continued usefulness of the racial-threat hypothesis" (p. 1168).

Using Voss's data and a slightly simplified version of his specification, I estimate the effect of percent black in the parish on white support for Duke in the same three elections by OLS and WLS.[18] The results are shown in table 3. In each election, the effect of percent black on white support for Duke is estimated to be at least twice as large by OLS as by WLS. In each case, percent black is statistically significant at conventional levels when estimated by OLS (with

[17] Louisiana's parishes are analogous to counties in other states. Votes for Duke among white cannot be directly observed. The author simply assume that no votes for Duke were cast by non-whites.

[18] Voss's complete specification includes an interaction term between percent black and percent urban which captures the possibility that racial threat may be a predominantly rural phenomena. The interaction term supports that notion when estimated by WLS, but is not statistically significant when estimated by OLS. Because it complicates the interpretation of results, I have left the interaction out of the analyses presented here.

Determinants of support for David Duke in LA Parishes
(OLS and WLS estimates)

	OLS			WLS		
	1990 Senate Primary	1991 Governor Primary	1991 Governor General	1990 Senate Primary	1991 Governor Primary	1991 Governor General
% Black	0.14	0.15	0.13	0.0085	0.08	0.04
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
Median white income	0.0004	-0.0002	0.0001	0.0002	-0.0006	0.0000
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
% Whites with H.S. educ.	-0.35	-0.24	-0.45	-0.38	0.11	-0.35
	(0.20)	(0.21)	(0.18)	(0.18)	(0.19)	(0.20)
% White unemployment	0.12	-0.33	0.09	0.07	-0.35	0.40
	(0.52)	(0.52)	(0.53)	(0.68)	(0.72)	(0.74)
% Urban	-0.08	-0.07	-0.07	-0.05	-0.09	-0.05
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
% White in-migration	-0.48	-0.70	-0.52	-0.55	-0.92	-0.60
	(0.28)	(0.33)	(0.22)	(0.17)	(0.18)	(0.18)
% French Speaking	-0.38	-0.72	-0.60	-0.49	-0.70	-0.61
	(0.09)	(0.10)	(0.08)	(0.09)	(0.09)	(0.09)
Constant	60.56	76.64	64.37	69.55	60.94	58.54
	(13.97)	(15.54)	(13.42)	(14.66)	(15.43)	(15.95)
R^2	0.65	0.75	0.70			
$\hat{\sigma}$	4.57	5.13	4.93			
N	62	62	62	62	62	62

Table 3: *The dependent variable is the parish Duke vote as a percentage of parish white voter registration. Data from Voss (1996), N=62. Standard errors in parentheses. Note that OLS support's Key's racial threat hypothesis while WLS does not.*

White's standard errors) and not significant when estimated by WLS. Thus, we have a case in which substantive conclusions depends upon the estimator selected. I will now turn to consider why this difference arises and which of these estimators in more likely to be producing accurate estimates.

If, like Voss, we assume that the election outcome in each parish is, in a sense, a random sample of the active electorate in each parish, we can estimate the standard error of that estimate as

$$\widehat{\text{SE}(P_i)} = \sqrt{\frac{P_i(1-P_i)}{n_i}}$$

where P_i is the percent support for Duke in a given election and n is the number of white registered voters in parish i. A histogram of these standard errors for the 1990 Senate election in shown in figure 4. The histogram highlights Voss's concern. Indeed there is a great disparity in the accuracy with which the underlying support for Duke is estimated between the large and small parishes. The standard error of the estimated support in the smallest parish (Tensas) is nearly 10 times as large as standard error of the largest parish (Jefferson). In the notation of the monte carlo results presented above, this is a case in which the variation in the sampling variances across observations, θ, is large and as such it is not surprising that it is a case in which OLS and WLS produce conflicting results.

While the variation in the sampling standard errors of the dependent variable is large, the absolute size of those standard errors are quite modest. On average, a 95 percent confidence interval around the estimated white support for Duke is only plus or minus one percentage point. Even the largest 95 percent confidence interval is only plus or minus 2.1 percentage points. Thus, overall, these estimates are quite reliable. A good gauge of whether WLS is appropriate in this case is the R^2 that the regression would achieve if the only source of error was the uncertainty due to estimation of the dependent variable. This R^2 can be approximated as one minus the average sampling variance of Y^* (the residual variance) over the observed variance of Y^* across parishes (the total variance).[19] In this case, the average estimated sampling error is about 0.24 while the inter-parish variance of white support for Duke in the 1990 Senate election is 57.9, so the R^2 that

[19]The approximate $R^2 = 1 - \frac{\bar{\omega}_i^2}{S_{y^*}}$ where S_{y^*} is the sample variance of the observed dependent variable across the observations on the dependent variable.

Distribution of standard errors of white support for David Duke across Parishes

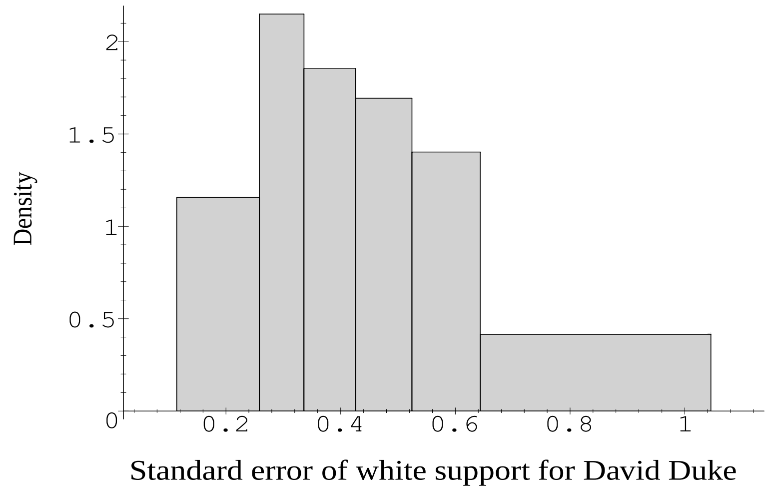

Figure 4: *Shows the distribution of standard errors of support for David Duke in the 1990 Louisiana Senate election across parishes.*

one would expect is greater than 0.99![20] Looking at table 3, we see that the R^2 from the OLS regression is a more modest 0.65. Thus, of the roughly 35 percent of the variance in the white support for Duke that is not accounted for by the model, only about 1 percent can be attributed to sampling errors in the measurement of the true underlying Duke support. In the notation of the monte carlo experiments, this would be a dataset in which C, the fraction of the error due to estimation of the dependent variable, was quite small. As such, we must conclude that the WLS estimates are unlikely to be accurate.

Table 4 estimates the same model of white support for Duke in the three elections using the two FGLS estimators presented above. The results confirm the conclusion that very little of the residual variation in the regression is due the estimation of the dependent variable. Both of the FGLS estimators produce estimates of the paramters that are nearly identical to the OLS estimates. Using the "known variance" estimator, we see that estimated σ (standard deviation of the homoscedastic component of the residual) was quite large in comparison to the average ω (standard error of the estimated the dependent variable). Therefore, the set of weights used in the second stage regression are dominated by σ and there is little variation in the weight from parish to parish. With so little

[20]That is, $0.995 = 1 - (0.24/57.9)$.

Determinants of support for David Duke in LA Parishes
(OLS and WLS estimates)

	Known variance estimator			Proportional variance estimator		
	1990 Senate Primary	1991 Governor Primary	1991 Governor General	1990 Senate Primary	1991 Governor Primary	1991 Governor General
% Black	0.14	0.15	0.13	0.15	0.18	0.11
	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
% Black	0.14	0.15	0.13	0.15	0.18	0.11
	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Median white income	0.0004	-0.0002	0.0001	0.0004	-0.0003	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
% Whites with H.S. educ.	-0.35	-0.23	-0.45	-0.35	-0.29	-0.46
	(0.17)	(0.19)	(0.18)	(0.17)	(0.19)	(0.19)
% White unemployment	0.12	-0.33	0.09	-0.07	-0.56	0.02
	(0.52)	(0.58)	(0.55)	(0.50)	(0.58)	(0.58)
% Urban	-0.08	-0.07	-0.07	-0.09	-0.05	-0.07
	(0.02)	(0.03)	(0.03)	(0.01)	(0.02)	(0.02)
% White in-migration	-0.48	-0.70	-0.52	-0.41	-0.67	-0.52
	(0.19)	(0.21)	(0.21)	(0.19)	(0.22)	(0.22)
% French Speaking	-0.38	-0.72	-0.60	-0.35	-0.75	-0.59
	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)
Constant	60.59	76.61	64.25	62.62	83.75	68.17
	(13.23)	(14.97)	(14.39)	(13.31)	(15.08)	(14.90)
$\hat{\sigma}$	4.54	5.11	4.91	5.02	5.58	5.24
Mean square ω	0.73	0.73	0.73			
Coef. on $1/n$				-74,558	-74,971	-67,350
				(26,439)	(32,748)	(31,966)
N	62	62	62	60	60	60

Table 4: *The dependent variable is the parish Duke vote as a percentage of parish white voter registration. Data from Voss (1996), N=62. Standard errors in parentheses. The "coef. on $1/n$" is the regression slope from regression of the squared OLS residuals on 1 over the number of registered voters in each precinct as described in the text. Note that both of these estimations support the racial threat hypothesis.*

variation in the weights, the similarity between the OLS and FGLS estimates come as no surprise. The "proportional variance" estimator yields very similar results. What is perhaps surprising about this example is that the overall heteroscedaticity in the regression residual is actually estimated to be *decreasing* in $1/n$.[21] Larger parishes appear to have significantly *higher* residual variances than do smaller parishes. Thus, there appears to be some evidence that in this case the component of the regression error that is not due to estimation of the dependent variable is heteroscedastic in a way that cancels and out and indeed swamps the heteroscedasticity induced by the estimation of the dependent variable.

This example reinforces the conclusions of the monte carlo experiments. In cases where only a small fraction of the total variation in the dependent variable is the result of sampling and where there is great heterogeneity in the sampling variance across observations, the traditional WLS approach can lead to misleading inferences. In these cases, OLS with robust standard errors and the FGLS approaches produce very similar results. Thus, while Voss was correct in noting that variation in parish size would lead to heteroskedasticity, his weighting approach appears to have done more harm than good. More appropriate estimators contradict Voss's conclusions and support the continued relevance of Key's racial-threat hypothesis.

5 Conclusion

The results presented in this paper suggest that information about the variance of the sampling errors in estimated dependent variables should be used with caution. The usual approach of weighting the dependent variable by of inverse of the standard errors of the dependent variable estimates will in most cases lead to inefficient parameter estimates and overconfidence in these estimates. This overconfidence can be very large. In some cases, WLS estimated parameter uncertainty was less than one-third as large as the true uncertainty. Discarding information about the sampling errors in the observations on the dependent variable and fitting OLS with White's robust standard errors is generally superior to the WLS approach. Indeed, OLS with White's standard errors is probably the best approach except when information about the sampling errors in the dependent

[21] The very large parameter value arise from the very small scale and range of $1/n$. For the 1990 Senate election $1/n$ ranges from 0.000006 to 0.0004.

variable is not only available, but highly reliable.

However, when reliable information about the sampling variances of the estimated dependent variable is available, the two FGLS approaches presented above will yield generally superior results to OLS. In some of the cases considered, the gain in efficiency was substantial (greater than 20 percent). These FGLS estimators are easy to implement and allow the analyst not only to achieve efficient parameter estimates, but also to estimate the fraction the total regression error that is due to sampling errors in the measurement of the dependent variable.

References

Beck, Nathaniel & Jonathan N. Katz. 1995. "What to do (and not to do) with Time-Series Cross-Section Data." *American Political Science Review*. 89(3):634–647.

Burden, Barry & David Kimball. 1998. "A New Approach to the Study of Ticket Splitting." *American Political Science Review*. 92(3):533–544.

Gay, Claudine. 1998. Taking Charge: Black Electoral Success and the Redefinition of American Politics PhD thesis Harvard University.

Gelman, Andrew & Gary King. 1994. "Enhancing Democracy Through Legislative Redistricting." *The American Political Science Review*. 88(3):541–559.

Giles, Michael W. & Melanie A. Buckner. 1993. "David Duke and Black Threat: An Old Hypothesis Revisited." *Journal of Politics*. 55(3):702–713.

Greene, William H. 1999. *Econometric Analysis*. Englewood NJ: Prentice Hall.

Hanushek, Eric A. 1974. "Efficient Estimators for Regressing Regression Coefficients." *The American Statistician*. 28(1):66–67.

Hanushek, Eric Alan & John E. Jackson. 1977. *Statistical methods for social scientists*. New York: Academic Press.

Herron, Michael & Kenneth Shotts. 2000. "Using Ecolgoical Inference Point Estimates in Second Stage Linear Regressions." Northwestern University Working Paper.

Key, V.O. 1949. *Suthern Politics*. New York: Knopf.

King, Gary. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton: Princeton University Press.

MacKuen, Michael B., Robert S. Erikson & James A. Stimson. 1989. "Macropartisanship." *American Political Science Review*. 83(4):1125–1142.

MacKuen, Michael B., Robert S. Erikson & James A. Stimson. 1992. "Peasants or Bankers? The American Electorate and the U.S. Economy." *American Political Science Review*. 86(3):597–611.

Maddala, G. S. 1992. *Introduction to Econometrics*. Englewood NJ: Prentice Hall.

Martinez, Michael D. 1990. "Partisan Reinforcement in Context and Cognition: Canadian Federal Partisanships, 1974–79." *American Journal of Political Science*. 34(3):822–845.

Oppenheimer, Bruce I. 1996. "The Representational Experience: The Effect of State Population on Senator-Constituency Linkages." *American Journal of Political Science*. 40(4):1280–1299.

Patel, Jagdish K. & Campbell B. Read. 1996. *Handbook of the normal distribution*. New York: Marcel Dekker.

Patterson, Samuel C. & Gregory A. Caldeira. 1985. "Mailing In the Vote: Correlates and Consequences of Absentee Voting." *American Journal of Political Science*. 29(4):766–788.

Roberts, Brian E. 2000. "The Redistributive Consequences of Changing Majority Control in the US House of Representatives." Presented the Midwest Political Science Association Annual Meetings, Chicago, April 2000.

Saxonhouse, Gary R. 1976. "Estimated Parameters as Dependent Variables." *American Economic Review*. 66(1):178–183.

Taylor, D. Garth. 1980. "Procedures for Evaluating Trends in Public Opinion." *Public Opinion Quarterly*. 44(1):86–100.

Voss, D. Stephen. 1996. "Beyond Racial Threat: Failure of the Old Hypothesis in the New South." *Journal of Politics*. 58(4):1156–70.

White, Halbert. 1980. "A Heteroscadastically-Consistent Covariance Matrix Estimator and a Direct Test for the Heteroscasticity." *Econometrica*. 48(4):817–838.