Combining Survey and Other Data To Estimate Agricultural Land Values

Charles H. Barnard and Danny Pfeffermann

Abstract *Combining survey, census, and administrative data improves the precision of survey estimates of mean agricultural land values A components-of-variance model is developed and applied to cropland value data for the Corn Belt Performance of the model compared with other procedures is tested using cross-validation techniques Results indicate that use of the proposed estimator would improve upon the USDA estimators at both the State and strata levels At the strata level particularly, the improvements may be very substantial*

Keywords. *Agricultural land values, components of variance, cross-validation, small areas, survey data*

The Economic Research Service (ERS) estimates farm real estate value for 48 States and the United States (*24*) [1] These estimates are derived from reports obtained in the Agricultural Land Values Survey (ALVS) from sampled farmers (*14,26*) The need for improved State-level values prompted an examination of alternative data sources and alternative estimators This paper shows how data available from sources other than the ALVS, and known at the county level, can be combined with ALVS data to improve the statistical precision of farmland value estimates The methods described here may be useful for improving the precision of other agricultural statistics

Small area (or small domain) estimation provides the foundation for a new estimator that combines data from the ALVS and other sources The basis for the new estimator is a prediction model that relates the individual farmers' reports to a set of regressor variables and a set of county and State effects The regressor variables measure known aggregate county characteristics, while the State and county effects represent specific influences not accounted for by the regressor variables In view of the large number of counties and the small sample sizes realized in all the counties and some of the States, the county and State effects are considered random giving rise to a mixed linear model with fixed regression coefficients and random components of variance The use of linear models with random effects is a common practice in small domain estimation The form of our model and the regressors included are chosen in order to best predict the small domain means of the target variable (the farmland values in the present case) and not necessarily to represent causal relationships with a substantive interpretation Indeed, while regression analysis has been used extensively to identify causal factors explaining the value of farmland, regression techniques have not been used to yield improved estimators (predictors) of mean farmland values

We show how the mixed linear model, in the context of small domain estimation, can potentially improve upon current USDA procedures Data from existing sources, measuring county characteristics that are believed to affect the farmland values, are selected as regressor variables Actual computation of the new estimator and its standard error (which we describe) permits an assessment of model performance and a comparison with USDA and other related estimators The results of that study conducted using cross-validation techniques, show that the new estimator in most instances substantially improves upon the estimators used by USDA, particularly at the strata level

Small Domain Estimation and the Mixed Linear Model

The problem underlying the computation of the farmland value indexes may be traced to the framework of survey sampling theory A survey population of all the farmers in the United States is divided into production regions, States, and counties The counties are grouped into homogeneous strata and a random sample of farmers is drawn from every stratum using a probability sampling plan [2] If the samples within the various strata were sufficiently large, one could

Barnard is an agricultural economist with the Agriculture and Rural Economy Division ERS Pfeffermann is senior lecturer Dept of Statistics Hebrew University Jerusalem The authors are grateful to Dr Ralph Folsom Research Triangle Institute NC for helpful discussions during early stages of the study and to Statistics Canada where Pfeffermann was a research fellow when the article was completed

[1]Italicized numbers in parentheses cite sources listed in the References section at the end of this article

[2]Strata consist in general of groups of adjacent counties called Crop Reporting Districts (CDR'S) Typically a State has 8-9 CRD's with each CRD consisting of 8-10 counties However urban-influenced counties have been extracted from CRD's and placed in special strata In what follows this definition of strata also defines the survey strata

estimate these means by the observed sample means, that is by averaging the farmland values reported by the farmers in the corresponding strata These strata estimates could then be averaged to produce State and regional estimates the usual USDA procedure However, the sample sizes selected by the ALVS within many States are too small to guarantee reliable State estimates, partly because of low response rates (The effective sample size in some States is less than 40) For instance, individual State indexes are not constructed for New England (27) On the other hand, for selected States where the estimates are more reliable, the National Agricultural Statistics Service (NASS) publishes indexes for strata Indexes for the United States and 10 major farm production regions may be considered reliable due to sufficiently large samples

The problem underlying the production of farmland indexes is a typical small-area estimation problem, an issue receiving increased attention in the statistical literature in recent years The problem of small-area estimation is that because of the small sample size in any given area the direct survey estimator based only on the sample observed for that area can become very unstable To overcome that problem, a variety of techniques has been proposed which essentially 'borrow strength" from one small area to the next, increasing the precision of the estimators in given small areas

The data used for these estimators include the observation on the target variable (the ALVS farmers' reports in our study) and the values of regressor variables x_1 x_k representing known small-area characteristics related to the unknown small-area means $\{\theta_i\}$ Denote by $\underset{\sim}{Y}_i$ the vector of observations on the target variable in small area i based on a sample of n_i units Assuming simple random sampling, it follows that $\overline{Y}_i = \sum_{j=1}^{n_i} Y_{ij} / n_i = \theta_i + \bar{\varepsilon}_i$ where $\bar{\varepsilon}_i$ is the corresponding mean of the error terms $\varepsilon_{ij} = Y_{ij} - \theta_i$, with expectation $E(\bar{\varepsilon}_i) = 0$ and variance $var(\bar{\varepsilon}_i) = \sigma^2_\varepsilon / n_i$

When the variances $\sigma^2_\varepsilon / n_i$ are suitably small, the statistician may be content to use the estimates $\overline{Y}_i$, which are basically the estimates currently used by USDA In view of the small sample sizes (n_i), however, other estimators have to be considered One natural candidate is the regression estimator $\underset{\sim}{x}'_i \underset{\sim}{b}$ where $\underset{\sim}{b}$ is a vector of estimated regression coefficients based on the individual observations Y_{ij} The estimator $\underset{\sim}{x}'_i \underset{\sim}{b}$ would be ideal if for every i, $\theta_i = \underset{\sim}{x}'_i \underset{\sim}{\beta}$ where $\underset{\sim}{\beta}$ represents the "true" unknown regression coefficients In fact, the estimator $\underset{\sim}{x}'_i \underset{\sim}{b}$ may still be used even when the relationship $\theta_i = \underset{\sim}{x}'_i \underset{\sim}{\beta}$ does not hold provided that the deviations $\{\theta_i - \underset{\sim}{x}'_i \underset{\sim}{\beta}\}$ are sufficiently small

Often the sample sizes n_i are too small to allow the use of the estimators $\{\overline{Y}_i\}$, and the deviations $\{\theta_i - \underset{\sim}{x}'_i \underset{\sim}{\beta}\}$ are too large to use the estimators $\underset{\sim}{x}'_i \underset{\sim}{b}$ Small-area estimation techniques are essentially a collection of models and inference procedures proposed in the literature to yield estimators that compromise between the estimators $\{\overline{Y}_i\}$ and $\underset{\sim}{x}_i \underset{\sim}{b}$, resulting in increased efficiency

For example, suppose it can be postulated that $\theta_i = \underset{\sim}{x}'_i \underset{\sim}{\beta} + v_i$, where $E(v_i) = 0$ and $var(v_i) = \sigma^2_V$ Notice that the deviations v_i are viewed now as random quantities Under this assumption, the model holding for the original observations can be written as a mixed linear model, $Y_{ij} = \underset{\sim}{x}'_i \underset{\sim}{\beta} + v_i + \varepsilon_{ij}$, where the individual error terms $\{v_i + \varepsilon_{ij}\}$ are now correlated within small areas due to the common effect v_i This model has been used by Battese, Harter, and Fuller (5) for the estimation of crop areas in counties in Iowa by using satellite data as the regressors The estimators derived under this model have the general form $\hat{\theta}_i = K_i \overline{Y}_i + (1-K_i)\, \underset{\sim}{x}'_i \underset{\sim}{b}$, where $K_i = \hat{\sigma}^2_V / \hat{\sigma}^2_V + \hat{\sigma}^2_\varepsilon / n_i$), and $\hat{\sigma}^2_V$ and $\hat{\sigma}^2_\varepsilon$ define suitable estimators of the unknown variances The estimator $\hat{\theta}_i$ is a weighted average of the estimators $\overline{Y}_i$ and $\underset{\sim}{x}'_i \underset{\sim}{b}$ with weights that reflect the relative precision of each of the two estimators In another article, Pfeffermann and Barnard (*17*) synthesize the recent research in small area estimation In what follows, we refer to that article using the abbreviation P-B

The model used in the present study extends the Battese-Harter-Fuller model by accounting for both State effects and nested county effects Let Y_{sci} be the farmland value reported in the ALVS by farmer i residing in county c of State s Let θ_{sc} stand for the unknown mean market value in county sc We postulate the following relationship

$$Y_{sci} = \theta_{sc} + \varepsilon_{sci}, \quad \theta_{sc} = \underset{\sim}{x}'_{sc} \underset{\sim}{\beta} + \alpha_s + \gamma_{sc}, \qquad (1)$$

for s=1 S, c=1 C(s), i=1 n_{sc},

where $\{\varepsilon_{sci}\}$ are independent errors with zero mean and variance σ^2_e, $\{\alpha_s\}$ represent random State effects with zero mean and variance σ^2_b, and $\{\gamma_{sc}\}$ are random county effects, nested within the State effects with zero mean and variance σ^2_w We assume that the three random components are mutually independent S is the total number of States in the study, C(s) is the total number of counties in State s, and n_{sc} is the number of reports in county c of State s

Equation 1 postulates that the land values reported by farmers residing in the same county, Y_{sci}, are distributed randomly around the true county mean, θ_{sc} The variation of the county means between counties is modeled as a function of known regressor variables, $\underset{\sim}{x}'_{sc}$, and random State and county effects The regressor variables represent k county characteristics with typical values represented by $\underset{\sim}{x}'_{sc} = (1, x_{sc1}, \ldots, x_{sck})$ for county sc (See the next section for the list of regressor variables used in our study)

The State effects represent any systematic influences on the prices of farmland that are common to all counties in a State, but that are not represented by the regressor variables State income and property tax laws, State environmental laws, and other regulatory policies that vary by State and restrict farm operation or landownership come to mind Similarly, the residual county effects represent unique county characteristics that systematically affect the values of farmland, but again, that are not represented by the regressor variables Examples might be the level of social services, school quality, and other characteristics that affect the quality of life

Substituting the right-side equation of (1) into the left-side equation gives the mixed linear (components of variance) model representation

$$Y_{sci} = \underset{\sim}{x}'_{sc}\underset{\sim}{\beta} + \alpha_s + \gamma_{sc} + \varepsilon_{sci}, \tag{2}$$

which implies

$$VAR(Y_{sci}) = \sigma_b^2 + \sigma_w^2 + \sigma_e^2,$$

$$COV(Y_{sci}, Y_{sci^*}) = \sigma_b^2 + \sigma_w^2, \; i \neq i^*,$$

$$COV(Y_{sci}, Y_{sc^*i^*}) = \sigma_b^2, \; c \neq c^*,$$

$$COV(Y_{sci}, Y_{s^*c^*i^*}) = 0, \; s \neq s^* \tag{3}$$

Thus, the model states that values reported by farmers residing in the same county are correlated, as are values reported by farmers residing in the same State but in different counties

The actual application of the model requires as a first step the identification of available data sources to be used as potential regressor variables We discuss this issue in the next section The formulas of the predictor of the county and State means, as obtained under the present model, are given later Note in this regard, that since the county and State means are considered as random under the model, we adopt hereafter the conventional statistical terminology and refer to the assessment of these means as "prediction" rather than "estimation "

Available Data. Sources and Definitions

Identification of factors that affect farmland values and statistical measurement of their importance has been the objective of studies for more than 60 years, spawning an extensive literature within the agricultural economics profession Reynolds (*18*), for instance, cites a partial list of more than 60 empirical studies The purpose of most of these studies was to discover the determinants of variation in farmland values and estimate the parameters associated with those factors Empirically, the general procedure is to regress observations on farmland values against corresponding observations on a set of independent variables representing variation in productivity or income, location relative to markets and services, and nonagricultural influences

In contrast, the purpose of our procedure is to identify regressor variables, which, when used jointly, can best predict the county and State mean farmland values without worrying about causal relationships and substantive interpretations Nevertheless, previous models fitted to farmland value data provide a natural basis to guide the selection of factors to include in our model The other obvious consideration in the preliminary selection of such regressor variables is data availability, which we describe next We then specify the variables considered in our study

Sources

The U S Department of Commerce (USDC) is a major source of county-level information that can be used as additional information to model the variation of farmland values The Census of Agriculture generally conducted every 5 years provides a wide array of agriculturally related information, including acres of land in farms, numbers of farms crop acres harvested, quantities of crops and livestock produced (sold), market values of crops and livestock sold, and days of off-farm work (*28*) Indeed, much of the literature involving cross-sectional analyses of aggregate farmland values has used county, State or national estimates provided by the Census of Agriculture In addition, the decennial Census of Population collects information on rural and urban population The Bureau of Economic Analysis, through its Regional Economic Information System, provides annual data on local area employment and personal income, by Standard Industrial Category (SIC) The data series available include mineral income, net farm income, and off-farm employment

A notable feature of the procedure presented in this article is its ability to include *alternative assessments of farmland value* among the regressor variables USDA, itself, collects farmland value information from three other independent sources, which provides direct assessments of county farmland values [3] Of particular interest is a set of data collected annually by the Agricultural Stabilization and Conservation Service (ASCS) from each of its County Executive Directors (CED's) This variable represents the opinions of ASCS county executive directors, one per county concerning the average value of nonirrigated cropland in their county Most CED's consult with farmers, lenders, and other real estate professionals before forming their opinions The data represent the opinion of each CED regarding average values of farmland in each county whose ASCS program they administer While there are some small differences in definition, these data provide an independent assessment of farmland values (*1,8*) and are particularly valuable because reports are received from virtually every agricultural county in the United States (See also the last two paragraphs at the end of this section)

Other sources of data on farmland values include the Farmland Market Survey, which obtains both sales data on individual tracts and opinion data on county farmland values (*27*) The opinion data from this survey are similar in form to the ASCS data but lack estimates for all counties The USDA Farm Costs and Returns Survey (FCRS) (*25*) and the Census of Agriculture (*3 28*) are sources of information on the value of farmland and buildings The Census of Agriculture provides data for every agricultural county but only at 5-year intervals Data from the FCRS, though available annually, are not available for every county

Variables Considered for Analysis

In studies cited by Reynolds (*18*) and Reynolds and Comer (*19*), many variables were found to be important determinants of farmland values Particular data used and specific results obtained have depended upon data availability and the level of aggregation employed Data used have varied from microdata on sales of individual tracts to aggregate data collected on a State, regional or national basis (*2*) In the selection and specification of variables for our model, we relied mostly on cross-sectional studies that used county data The variables chosen are general in the sense that they could be used in the analysis of nonirrigated cropland in most regions of the United States Our initial model included 10 variables eight that represent various aspects of agricultural productivity and urban influence, one that represents mining activity, and one that is the independent assessment of nonirrigated cropland value from the CED's of ASCS Brief descriptions of the variables selected, their sources, and abbreviated names are provided in table 1, with more detailed explanations following

Various measures have been used to represent agricultural productivity and the overall economic potential of farmland (*4,6,9,15,16,21,22*) In our model, PCTFARM and PCTGRAZ were included in the model to represent the basic suitability of the land for crops, which depends on soil, climate, topography, and other factors A larger percentage of farmland generally indicates higher average productivity, while larger percentages of grazing land indicate lower average productivity Variation in overall economic potential of land for agricultural use is measured by FARMINCOME and CROPSVALU Larger net farm incomes and gross crop receipts per acre imply more productive cropland Although these variables are crude measures individually, taken together they serve as proxies for the agricultural value of farmland

A similar variable, SPECLTYVALU, was included to capture the contribution of high-valued specialty crops, including vegetables, fruits, berries, nuts, and greenhouse products Such variables measure differences in land use intensity The importance of specialty crops as a determinant of land values is demonstrated by Reynolds and Tseng (*21*) in a study of Florida counties

Size of tract has been demonstrated to be an important explanatory variable in models designed to explain farmland values (*6,7 12,15,16,21,22,23*) Our variable, FARMSIZE, serves as a proxy for size of tract sold Value per acre declines as tract size increases, *ceteris paribus*

Measures of urban influence, including population, population density, and extent of off-farm employment, have often been found to have large and statistically significant effects on farmland value (*6,7,9, 12,13,15,16 20,21,22,23*) In our study, these nonagricultural influences are represented by POPULATN and NUMOFFFARM Larger urban populations imply increased demand for farmland for rural residences More off-farm employment opportunities imply increased potential for part-time and hobby farms Nonagricultural uses often can outbid agriculture for use of urban-influenced farmland

MINEINCOME is another variable related to nonagricultural influences included principally to capture the effect that mineral rights may have on the sale price of individual parcels When opinions of value are formed on the basis of reported farmland sale prices, a portion of the value of mineral rights

[3]All four USDA sources collect data during January-February of each year

Table 1—Variables used in the empirical study

Abbreviated name	Description	Source
CED	County executive directors' opinions of mean county farmland values	ASCS questionnaire[1]
PCTFARM	Acres of farmland as a percentage of county land area	Census of Agriculture[2]
POPULATN	Urban population per acre of total cropland	Census of Population[2] Census of Agriculture[2]
CROPSVALU	Market value of crops sold per acre of total cropland	Census of Agriculture
FARMINCOME	Net farm income per acre of land in farm	Local area personal income[3]
FARMSIZE	Average number of acres per farm	Census of Agriculture
PCTGRAZ	Acres of grazing land as a percentage of land in farms	Census of Agriculture
SPECLTYVALU	Market value of specialty crops per acre of total cropland	Census of Agriculture
NUMOFFFARM	Number of farm operators who worked at least some days off the farm	Local area employment[3]
MINEINCOME	County income from mining	Local area personal income[3]

[1]Agricultural Stabilization and Conservation Service USDA
[2]Bureau of the Census USDC
[3]Bureau of Economic Analysis USDC

may be incorporated This factor may positively affect cropland values in areas with substantial oil, gas, and coal development

The final variable included in our initial model was the average value of nonirrigated cropland from ASCS described earlier, labeled CED in table 1 The inclusion of the CED variable as one of the regressor variables raises two interesting questions The first question refers to the different roles assigned to the CED and the ALVS measurements, namely, one variable being specified as an independent variable and the other as the dependent variable despite the fact that both variables measure essentially the same phenomenon Our consideration in including the CED variable as the regressor variable was that this variable, unlike the ALVS, is measured in every county and can be used in the model without missing observations Also in contrast to the ALVS estimates, whose precision depends on the realized sample sizes, which differ from one county to the other, the CED variable uses the same sort of information in every county Theoretically a better way to include these variables in the model would have been to specify both of them as dependent correlated variables Notice, however, that this multivariate framework is much more complicated computationally, whereas the gains in terms of the efficiency of the resulting predictors would generally be low considering that both the univariate and the multivariate models exploit the same amount of information If the joint distribution of the two estimators can be assumed to be bivariate normal then the structure of the predictors as obtained under the two models is similar (even though not the same)

The other question applies to the interpretability of the model In some sense, the CED variable encompasses and measures the interaction of all the other regressor variables included in the model and as such, the model has no longer a substantive causal interpretation We re-emphasize however, that the purpose of the analysis is the prediction of the county and State means Thus, variables have been included in the model based on their prediction power and not with respect to their substantive interpretation, an important factor when analyzing the results of this study

Computation of the Predictors and Prediction MSE's

In this section, we outline the major stages in fitting the model defined by equations 2 and 3 to the actual data A more technical and comprehensive discussion can be found in the P-B article We assume a given set of regressor variables with typical values $\underset{\sim}{X}'_{sc} = (1, x_{sc1} \ldots x_{sck})$ corresponding to county sc

Presentation of the Model in Matrix Notation

Let $\underset{\sim}{Y}_{sc}$ represent the vector of observed values in county sc, and let $\underset{\sim}{Y}'_s = (\underset{\sim}{Y}'_{s1} \ldots \underset{\sim}{Y}'_{sC(s)})$ define the vector of observations in State s so that $\underset{\sim}{Y}' = (\underset{\sim}{Y}'_1 \ldots \underset{\sim}{Y}'_s)$ defines the entire vector of reported land values A similar notation is used for the residuals $\{\varepsilon_{sci}\}$ We denote by $\underset{\sim}{\alpha}' = (\alpha_1 \ldots \alpha_s)$ the vector of State effects and by $\underset{\sim}{\gamma}' = (\gamma_{11} \ldots \gamma_{1C(1)}, \ldots, \gamma_{S1} \ldots \gamma_{SC(S)})$ the vector of nested county effects of order $T_\gamma = \sum_{s=1}^{S} C(s)$

Using the symbol $\otimes$ to define the Kronecker product, $n_s = \sum_{c=1}^{C(s)} n_{sc}$ to represent the number of observations in State s, and $\underset{\sim}{1}'_m$ to define in general a $1 * m$ vector of ones, the model defined by equation 2 can be written compactly as

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + Z_b\underset{\sim}{\alpha} + Z_w\underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} = X\underset{\sim}{\beta} + \underset{\sim}{\mu}, \tag{4}$$

where $X' =$

$$[\underset{\sim}{1}'_{n_{11}} \otimes \underset{\sim}{x}_{11}, \underset{\sim}{1}'_{n_{12}} \otimes \underset{\sim}{x}_{12}, \ldots, \underset{\sim}{1}_{n_{1C(1)}} \otimes \underset{\sim}{x}_{1C(1)} \ldots \underset{\sim}{1}'_{n_{S1}} \otimes \underset{\sim}{x}_{S1}, \ldots, \underset{\sim}{1}'_{n_{SC(S)}} \otimes \underset{\sim}{x}_{SC(S)}]$$

$$Z_b = \begin{bmatrix} \underset{\sim}{1}_{n_1} & & \\ & \ddots & \\ & & \underset{\sim}{1}_{n_S} \end{bmatrix}$$

$$Z_w = \begin{bmatrix} \underset{\sim}{1}_{n_{11}} & & \\ & \ddots & \\ & & \underset{\sim}{1}_{n_{SC(S)}} \end{bmatrix}$$

The vector $\underset{\sim}{\mu}$ satisfies

$$E(\underset{\sim}{\mu}) = \underset{\sim}{0},$$

$$E(\underset{\sim}{\mu}\underset{\sim}{\mu}') = \sigma_b^2 Z_b Z'_b + \sigma_W^2 Z_W Z'_W + \sigma_e^2 I_n, \tag{5}$$

where $n = \sum_{s=1}^{S} n_s$ and I_n is the identity matrix of order n

Optimal Predictors of County and Strata Means Assuming Known Variances

The optimal predictors of the county and Strata means are obtained in a straightforward manner from the optimal predictor of the vector $\underset{\sim}{\lambda}' = (\underset{\sim}{\beta}', \underset{\sim}{\alpha}', \underset{\sim}{\gamma}')$ One way to derive the optimal predictor $\hat{\underset{\sim}{\lambda}}$ and the associated variance-covariance (V-C) matrix of the prediction errors $(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda})$ is to compute $\hat{\underset{\sim}{\lambda}}$ as the generalized least squares (GLS) solution of the regression model

$$\underset{\sim}{Y}^0 = \begin{bmatrix} \underset{\sim}{Y} \\ \\ \underset{\sim}{0}_r \end{bmatrix} = \begin{bmatrix} X & Z_b, & Z_w \\ \\ {}_r0_{k+1}, & & -I_r \end{bmatrix} \begin{bmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\alpha} \\ \underset{\sim}{\gamma} \end{bmatrix} + \begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\alpha} \\ \underset{\sim}{\gamma} \end{bmatrix} = X^0\underset{\sim}{\lambda} + \underset{\sim}{\varepsilon}^0, \tag{6}$$

where $\underset{\sim}{0}_r$ and ${}_r0_{k+1}$ define correspondingly a column null vector of order $r = S+T_\gamma$ and a null matrix of order $r * (k+1)$ (*10*) The error vector $\underset{\sim}{\varepsilon}^0$ satisfies $E(\underset{\sim}{\varepsilon}^0) = \underset{\sim}{0}$, $E(\underset{\sim}{\varepsilon}^0\underset{\sim}{\varepsilon}^{0\prime}) = V = \mathrm{Diag}[\sigma_e^2\underset{\sim}{1}'_n, \sigma_b^2\underset{\sim}{1}'_S, \sigma_w^2\underset{\sim}{1}'_{T_\gamma}]$ The GLS estimator of $\underset{\sim}{\lambda}$ is

$$\hat{\underset{\sim}{\lambda}} = (X^{0\prime}V^{-1}X^0)^{-1} X^{0\prime}V^{-1}\underset{\sim}{Y}^0 \tag{7}$$

Notice that X^o is of full rank (assuming X is of full rank), which guarantees a unique solution The V-C matrix of the prediction errors has the common form

$$\mathrm{VAR}(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda}) = \mathrm{E}\,(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda})(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda})' = (X^{o\prime}V^{-1}X^{o})^{-1}, \qquad (8)$$

where the block matrix consisting of the first (k+1) rows and columns of $(X^{o\prime}V^{-1}X^{o})^{-1}$ is the V-C matrix of the GLS estimator $\hat{\underset{\sim}{\beta}}$ of $\underset{\sim}{\beta}$ [4]

The optimal predictors of the county means and the corresponding prediction variances are obtained from $\hat{\underset{\sim}{\lambda}}$ and $\mathrm{VAR}(\hat{\underset{\sim}{\lambda}} - \underset{\sim}{\lambda})$ as

$$\hat{\theta}_{sc} = \underset{\sim}{x}'_{sc}\hat{\underset{\sim}{\beta}} + \hat{\alpha}_s + \hat{\gamma}_{sc} = \underset{\sim}{h}'_{sc}\hat{\underset{\sim}{\lambda}}, \qquad (9)$$

$$\mathrm{E}(\hat{\theta}_{sc} - \theta_{sc})^2 = \underset{\sim}{h}'_{sc}\,(X^{o\prime}V^{-1}X^{o})^{-1}\underset{\sim}{h}_{sc}, \qquad (10)$$

where $\underset{\sim}{h}'_{sc} = (\underset{\sim}{x}'_{sc}, \underset{\sim}{q}'_{sc})$ and $\underset{\sim}{q}'_{sc}$ is a row vector of length $(S+T_\gamma)$ with 1's in positions s and $(\sum_{t=0}^{s-1} C(t) + c)$ and zeros elsewhere, $[C(0) \equiv 0]$

The mean farmland values of the survey strata are obtained from the county means as

$$\theta_{sh} = \sum_{sc \in sh} a_{sc}\theta_{sc} \Big/ \sum_{sc \in sh} a_{sc} = \sum_{sc \in sh} \tilde{a}_{sc}\theta_{sc}\,, \qquad (11)$$

where a_{sc} is the total acreage of the particular type of farmland in county sc, $\tilde{a}_{sc}$ is the proportion of acreage of that particular type of farmland in stratum sh that is found in county sc, and the summation is over counties in State s belonging to stratum h Since θ_{sh} is a linear combination of the county means, it follows that the optimal predictors of the survey strata means are

$$\hat{\underset{\sim}{\theta}}_{sh} = \sum_{sc \in sh} \tilde{a}_{sc}\hat{\theta}_{sc} = \hat{\alpha}_s + \sum_{sc \in sh} \tilde{a}_{sc}\,(\underset{\sim}{x}'_{sc}\hat{\underset{\sim}{\beta}} + \hat{\gamma}_{sc}) = \underset{\sim}{f}'_{sh}\hat{\underset{\sim}{\lambda}} \qquad (12)$$

where $\underset{\sim}{f}'_{sh} = (\sum_{sc \in sh} \tilde{a}_{sc}\,\underset{\sim}{x}'_{sc}, \underset{\sim}{r}'_{sh})$ and $\underset{\sim}{r}'_{sh}$ is a row vector of length $(S + T_\gamma)$ with one in position s, $\tilde{a}_{sh1}$ $\tilde{a}_{shC(sh)}$ in the positions corresponding to counties included in stratum sh, and zeros elsewhere C(sh) is the number of counties included in stratum h of State s For example, s=1, if S=5, and h=1, then $r'_{11} = (1,0,0,0,0,|\,\tilde{a}_{111} \quad \tilde{a}_{1C(11)},0, \quad 0\,|\,0, \quad 0)$ The prediction variance of $\hat{\theta}_{sh}$ is

$$\mathrm{E}(\hat{\theta}_{sh} - \theta_{sh})^2 = \underset{\sim}{f}'_{sh}\,(X^{o\prime}V^{-1}X^{o})^{-1}\underset{\sim}{f}_{sh} \qquad (13)$$

The use of equations 9-13 assumes that the sample includes farmers from every county P-B gives the appropriate formulas for the case where some of the counties are not represented in the sample The optimal predictors of the State means can be obtained in similar fashion

Variance Estimation

The discussion to this point assumes known variances In practice, the variances have to be estimated from the sample P-B discusses the practical aspects of estimating the unknown variances by maximum likelihood methods assuming that the model random disturbances have a normal distribution They illustrate that the variance estimates can be obtained by iterating between the procedures "REG" and "VARCOMP" in SAS

Substituting the sample estimates for the true variances in the formulas for $\hat{\theta}_{sc}$ and $\hat{\theta}_{sh}$ gives the corresponding empirical predictors of the county and strata means Performing a similar substitution in the formulas of the V-C matrices yields, in the case of large samples, the V-C matrices of the empirical predictors These matrices have to be modified in the case of small sample sizes in order to account for the extra variability induced by the need to estimate the unknown variances See, for example, Kackar and Harville (*11*)

Application of the Model

The model defined in the previous section was applied to data collected by the ALVS The purpose of this analysis was twofold to test the suitability of the model to the land values data, and to compare the performance of the model-dependent predictors with the performance of other possible predictors (estimators), including the survey estimator used by USDA

The USDA Survey Estimator

The ALVS is an opinion survey of farmers and ranchers Participants in the survey are selected by a stratified simple random design, carried out separately within each of the States, with a 20-percent sample rotation from one year to the next The questionnaire asks for information on average market value per acre of irrigated and nonirrigated cropland, grazing land, and woodland The values reported by the farmers are averaged first within strata and then over the strata within States to yield estimates of

[4] An important advantage of expressing $\hat{\lambda}$ as the GLS solution of the regression model (equation 6) is that the predictor and the prediction V-C matrix can be computed using any computer software for weighted regression with Y^0 as the dependent variable X^0 as the design matrix, and $\underset{\sim}{w}' = (\underset{\sim}{1}'_{n+r})\,V^{-1}$ as the vector of weights

State average market value, by type of farmland (*14*) Until 1989, the averages within strata were simple means, while the averages of the strata means were weighted averages, the weights being relative to the total acreage of the particular type of farmland in the given strata ERS changed its procedure in 1989, and the strata estimates are now weighted averages of county means Acreages come directly from, or are derived from, the latest Census of Agriculture (*28*)

Application of the Model to Corn Belt Data

The survey data analyzed in this study are the values of nonirrigated cropland in the Corn Belt States as collected in the 1984 ALVS [5] Nonirrigated cropland constitutes the major land use in the region The data consist of 871 farmers' reports representing 5 States (Indiana, Illinois, Iowa, Missouri, and Ohio), 43 strata, and 251 counties We excluded from the analysis the strata formed for the urban-influenced counties (see footnote 2) since the farmland values in these strata behave very differently from the values in the other strata, thus requiring extra treatment [6] In urban-influenced counties, particularly those that are part of large metropolitan areas, farmland values are higher and have larger variances than counties in more rural areas The mean and variances of farmland values in the excluded strata are 38 percent and 339 percent higher than in the remaining strata, respectively Farmland values in urban-influenced counties exhibit little relationship to the agricultural characteristics that determine farmland values in rural counties, suggesting the need for alternative model specification Although the current model does contain a proxy for urban influence (the POPULATN variable), county-level population cannot fully account for the influence of large multicounty metropolitan areas Distance from the center of the county to the center of the nearest major metropolitan area might more accurately account for the variation in the excluded strata Distance measures have been used in previous studies with good success Such measures are not available from published sources, but future work should involve the development of such data

The 10 variables listed in table 1, plus an added intercept, formed the initial X matrix for the model (equation 4), while the dollar per acre values reported in the ALVS constituted the Y vector The model was estimated based on the entire data set The significance of the β coefficients was tested by using the Wald statistic (*29*) The six variables listed in the lower part of the table and the intercept variable were jointly insignificant in the presence of the other four variables (As discussed before, the emphasis in the present study is on prediction rather than on interpretation, so we chose to include variables with significant predictive power rather than variables necessarily having substantive interpretation) Consequently, the nonsignificant variables were excluded from the model and were not considered in the rest of the analysis (The Wald statistic for testing a hypothesis of the form H_0 $C\beta = 0$, where C is $r * (k+1)$, is $W = (C\hat{\beta})'$ $[C\ VAR(\hat{\beta})\ C']^{-1}C\hat{\beta}$, and it has an asymptotic chi-square distribution with r degrees of freedom under H_0 The value observed when testing the joint significance of the seven variables was W=4 15, which was well below the customary critical values of that $\chi^2_{(7)}$ distribution)

Table 2 shows the four significant regression coefficients (first four elements of the empirical predictor $\hat{\lambda}$) along with their estimated standard errors, the variance components estimates, and twice the log of the likelihood ratio test statistic (log LRT) used for testing significance These test values indicate highly significant variance component estimates as can be seen by comparing the test values to critical values of the $\chi^2_{(1)}$ distribution The test results should be interpreted with caution, since the postulated chi-square distribution is a large sample property, whereas the data represent only five groups

Table 2 reveals the highly significant nature of the CED variable, which is by far the most important predictive variable To illustrate the importance of this variable, we conducted the following simple analysis, using ordinary least squares regression (OLS) An equation containing only the CED variable and an intercept was estimated and compared with an equation containing the four significant variables and an intercept The regression sum of squares for the CED-only equation amounted to 96 percent of the regression sum of squares for the latter equation Dropping the CED variable and estimating an equation containing only an intercept and the other three significant variables results in a 30-percent reduction in the regression sum of squares

The dominant predictive power of the CED variable (available from ASCS data) is especially important because the information it contains is updated annually and in the same time period as the ALVS This contrasts with the 5-year periodicity of information from the Census of Agriculture

[5] Restriction of the analysis to the Corn Belt was mainly for technical reasons, but this region, nevertheless, sufficiently illustrates the important features of the proposed procedures

[6] The Corn Belt consists of 495 counties 49 are part of the excluded urban-influenced strata and 195 had no observations in the ALVS

Table 2—Significant regression coefficients and variance components

Item	CED	PCTFARM	POPULATN	CROPSVALU
Regression coefficients	0 59	663 8	357 4	811 1
Standard errors	05	118 9	148 1	498 3

Variance components	Significance tests	2 logLRT
Between States—σ_b^2 = 24,337	$H_0\ \sigma_b^2 = 0$	60 2
Between counties—σ_w^2 = 24,157	$H_0\ \sigma_w^2 = 0$	31 6
Residual—$\sigma\pi_e^2$ = 174,940		

Testing the Performance of the Model

To assess the performance of the model in predicting the unknown strata and State means, we performed a cross-validation study by which the model-based predictor and other estimators were calculated based on one part of the sample (the estimation part) The performance of the predictor and estimators has been evaluated based on their quality in predicting the data included in the complementary part (the validation part) This method differs from the direct analysis of all the data reported in the P-B article, with the advantage that the assessment and comparison of the various estimators and predictors are less tied to a particular model The results obtained from the study, however refer to the sample sizes of the partitioned data sets and not to the sample sizes of the combined sample, which are the actual sample sizes of the ALVS

We split the sample between counties within strata About half the counties of each stratum were allocated to the estimation part and the other half to the validation part We employed a simple random sampling design for the splitting algorithm

We evaluated the performance of four predictors of the survey strata means by computing the prediction bias and root mean square error (RMSE) of the predictors and averaging the results within States by using the relative strata acreages as weights The strata-based analysis enables a comparison with the survey estimator used until 1989, which is defined by USDA as an unweighted average at the strata level [7] Thus, let $\tilde{M}_{sh}$ represent any one of the four predictors and $M_{sh} = \sum_{sc \in V_{sh}} a_{sc} \overline{Y}_{sc} / \sum_{sc \in V_{sh}} a_{sc}$ define the mean for farmers included in the validation part of stratum h in State s, where a_{sc} is the acreage of nonirrigated cropland in county sc and $\overline{Y}_{sc}$ is the sample mean of observations in county sc As such, the prediction BIAS and RMSE are represented by

$$\mathrm{BIAS}_s(\tilde{M}_{sh}) = \left(\sum_h a_{sh}(\tilde{M}_{sh} - M_{sh})\right) / \sum_h a_{sh}, \qquad (14)$$

and

$$\mathrm{RMSE}_s(\tilde{M}_{sh}) = \left(\sum_h a_{sh}(\tilde{M}_{sh} - M_{sh})^2 / \sum_h a_{sh}\right)^{1/2}, \qquad (15)$$

where a_{sh} is the acreage of nonirrigated cropland in stratum sh and the summation $\sum_h$ is over all the strata included in State s

Using the prediction bias (equation 14) and RMSE (equation 15) as criteria, we compare the performance of the following predictors of the strata means

A The USDA survey estimates, $\dot{M}_{sh}$, which were defined as

$$\dot{M}_{sh} = \sum_{sc \in E_{sh}} n_{sc} \overline{Y}_{sc} / \sum_{sc \in E_{sh}} n_{sc}, \qquad (16)$$

where the summation is over counties from stratum h in State s included in the estimation part

[7] Our cross-validation study was initially designed to evaluate the pre-1989 USDA estimator Since we are trying to predict strata means over counties included only in the validation part there was no apparent reason to prefer the new USDA estimator over the old USDA estimator Supplemental analysis indicates that a comparison between our estimator and either of the two USDA estimators is essentially independent of the weighting procedure used In the augmented analysis, we considered a second split which allocated approximately half of the farmers of each county to the estimation part and the rest of the farmers to the validation part To reflect more closely the new procedure used by USDA in 1989, we weighted our county predictions by county acreages The results obtained for that second split are generally consistent with results reported here

B The optimal predictors, $(\hat{\theta}^{E}_{sh})$, where the superscript "E" added to the previous notation is used to emphasize that the predictors have been calculated based on the estimation part and that the unknown variances have been replaced by the sample estimates The optimal strata means predictors are defined as

$$\hat{\theta}^{E}_{sc} = \sum_{sc \in V_{sh}} a_{sc}\,\hat{\theta}^{E}_{sc} / \sum_{sc \in V_{sh}} a_{sc} \tag{17}$$

where

$$\hat{\theta}^{E}_{sc} = \underset{\sim}{x}'_{sc}\underset{\sim}{\hat{\beta}} + \hat{\alpha}_{s} \tag{18}$$

The county effects, $\hat{\gamma}_{sc}$, are estimated as zero because the sample was split between counties, so that counties sc selected for the validation part are not represented in the sample

C The synthetic regression estimators, $\hat{R}_{sh}$, which are calculated as weighted averages of the county regression estimators, $\hat{R}_{sc} = \underset{\sim}{x}'_{sc}\underset{\sim}{\hat{\beta}}$, where $\underset{\sim}{\hat{\beta}}$ is the optimal maximum likelihood estimator (mle) of $\underset{\sim}{\beta}$ and the weighting procedure used is the same as that defined above for the optimal estimators

D The synthetic regression estimators, $\hat{R}^{ols}_{sh}$, which are calculated in the same way as the estimator $\hat{R}_{sh}$ except that $\underset{\sim}{\beta}$ is estimated using ordinary least squares

The synthetic regression estimators, $\hat{R}_{sh}$ and $\hat{R}^{ols}_{sh}$, represent alternative estimators that also incorporate the county-specific information The estimator $\hat{R}_{sh}$ accounts for the correlations between the various farmers' opinions which result from the common county and State effects (see equation 3) Specific estimates of the State and county effects, however, are not incorporated into this estimator The OLS estimator, on the other hand, ignores State and county effects altogether

Table 3 gives the prediction bias and RMSE of the various predictors separately for each State Also shown are the target weighted averages of the strata means in the validation part defined as $M_s = \sum_h a_{sh}M_{sh} / \sum_h a_{sh}$ [8]

[8] A robust predictor, incorporating a restriction to assure that the mean farmland value predicted under the model for the entire group of States will equal the survey estimator of that same mean, is derived in P-B The bias and RMSE of the robust predictor came out very similar to those of the optimal predictor This outcome can be considered indicative of the adequacy of the model

The main conclusion to note from the table is that the use of the alternative data sources improves the prediction of the farmland values The improvement is evident at the State level as revealed by comparing the prediction biases of the optimal predictor and the USDA survey estimator The prediction bias of the optimal predictor is substantially lower in four of the five States Among the three predictors using the additional information, the optimal predictor is clearly the most accurate, demonstrating the benefit of accounting for State and county effects in the form of a variance components model The two synthetic regression estimators show improvement relative to the USDA survey estimator in the prediction of the State means in two States, but the estimators actually perform less well than the USDA survey estimators in the other three States, particularly in Missouri where they miss by a wide margin

The RMSE of the optimal estimator is lower than the RMSE of the USDA survey estimator in three States The reduction amounts to approximately 50 percent in two of those States The RMSE of the optimal estimator in the remaining two States is only slightly larger than for the USDA survey estimator The two synthetic estimators also show a reduction in the RMSE relative to the USDA survey estimator in three States, but the reduction is less pronounced than for the optimal estimator For Missouri, the RMSE's of the synthetic estimators are considerably larger than the RMSE of the USDA survey estimator.

The use of the additional information not only improves upon the USDA survey estimators in terms of point predictions but also provides a basis for probabilistic inference Table 4 contains the 95-percent prediction intervals for the validation State means The prediction intervals are of the form $\hat{\theta}^{E}_{s} \pm Z_{\alpha/2}[\widehat{VAR}(\hat{\theta}^{E}_{s} - M_s)]^{1/2}$ where the model-dependent estimates of the prediction variances are used in the calculation The notable result from table 4 is that the validation mean is within the prediction interval The bias $(\hat{\theta}^{E}_{s} - M_s)$ is less than $1\ 96[\widehat{VAR}(\hat{\theta}^{E}_{s} - M_s)]^{1/2}$ in all five States, indicating the insignificance of the prediction bias at the 5-percent level

Conclusions and Model Extension

The results of the empirical study indicate that the use of alternative data sources improves the precision of mean farmland value estimates Consideration of

Table 3—Bias and root mean square errors of strata means predictors

Item	Criteria	Indiana	Illinois	Iowa	Missouri	Ohio
				Dollars per acre		
State means (validation)		1,958	1,689	1,663	822	1 484
Predictors						
$\hat{M}_{sh}$	BIAS	-116 8	49 9	73 5	-25 5	-114 5
	RMSE	378 7	278 0	190 5	128 7	331 7
$\hat{\theta}^{E}_{sh}$	BIAS	-31 4	-4 1	23 5	83 2	-67 5
	RMSE	182 6	214 5	199 0	141 3	184 1
$\bar{R}_{sh}$	BIAS	-187 5	-93 0	-3 1	268 3	19 0
	RMSE	259 8	233 8	197 6	291 6	172 3
$\hat{R}^{ols}_{sh}$	BIAS	-137 2	-60 1	21 4	202 9	19 9
	RMSE	220 4	242 0	207 5	235 1	180 3

Table 4—Confidence intervals for the validation State means

State	Upper limit	Validation mean	Lower limit
		Dollars per acre	
Indiana	2,073	1,958	1,779
Illinois	1,879	1,689	1,507
Iowa	1,815	1,663	1,557
Missouri	1,074	822	736
Ohio	1,580	1,484	1,254

State and county effects in the form of a nested, variance-components model adds to the precision of the assessments The computations involved in the application of the procedure can be performed using available statistical software In addition, the model provides a satisfactory basis for probabilistic inference

Although the study demonstrated the potential for predictors derived under the model to improve substantially upon the estimators used by USDA, the results strictly apply only to a major land use (nonirrigated cropland) in a very homogeneous farm production region (the Corn Belt) The procedure's ability to produce improvements for irrigated cropland, grazing land, and woodland in more heterogeneous regions is yet to be tested A full evaluation would also involve extension of the model to include more States in the analysis and the consideration of additional regressor variables The inclusion of more States will provide more stable estimates for the variance components and, hence, better predictors of State and strata farmland values Notice in this respect that it is unnecessary to assume the same regression coefficients for all regions By appropriate definition of the design X-matrix, different vectors of coefficients can be postulated for different regions

Consideration of additional regressor variables may improve the predictions Variables that jointly account for both population of major metropolitan areas and county location relative to those areas may be especially helpful Such variables, which represent access to social services, recreational facilities, and other quality-of-life conditions, may be most useful in modeling farmland values in the urban-influenced strata that were excluded from this study

As a final note, we point out the potential applicability of this procedure to a wide variety of data obtained from surveys conducted by ERS and NASS With appropriate modification, the procedures could be applied, for example, to farmland value data obtained in the FCRS

References

1 Barnard, C H *U S Farmland Values, April 1982 An Experimental Survey of ASCS County Executive Directors* Staff Report No AGES830706, U S Dept Agr , Econ Res Serv , Sept 1983

2 Barnard, C H "Data Sources for Farm Real Estate Market Studies," *The Farm Real Estate Market* SRDC Publication No 83, Mississippi State Univ , 1986, pp 15-28

3 Barnard, C H , and J Jones *Farm Real Estate Values in the United States by Counties, 1850-1982* SB-751, U S Dept Agr , Econ Res Serv , Mar 1987

4 Barrows, R L , and R W Dunford *Agricultural, Urban, and Recreational Determinants of Farmland Values in Wisconsin* University of Wisconsin, Madison, College of Agricultural and Life Science, Research Bulletin R2764, 1976

5 Battese, G E , R M Harter and W A Fuller "An Error Components Model for Prediction of County Crop Acres Using Survey and Satellite Data," *Journal of the American Statistical Association* Vol 83, 1988, pp 28-36

6 Clifton, I D "Using Automatic Interaction Detection (AID) Model to Obtain Homogeneous Classifications of Farmland Markets," *Agricultural Economics Research* Vol 27 1975, pp 93-100

7 Greene, C , and C H Barnard "Agricultural Land Markets in U S Urban-Fringe Counties," *Landscape and Urban Planning* Vol 16, 1988, pp 293-301

8 Greene, C , and C H Barnard *U S Farmland Values, 1982-84 A Comparison of Experimental and Traditional Data* Staff Report No AGES850304, U S Dept Agr , Econ Res Serv , July 1985

9 Hammill, A E "Variables Related to Farm Real Estate Values in Minnesota Counties," *Agricultural Economics Research*, Vol 21, 1969

10 Harville, D A "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of the American Statistical Association* Vol 72, 1977, pp 320-40

11 Kackar R N , and D A Harville "Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models," *Journal of the American Statistical Association* Vol 79, 1984, pp 853-62

12 Morris, D E "Farmland Values and Urbanization," *Agricultural Economics Research* Vol 30, 1978, pp 44-47

13 Norberg, R P , and J E Reynolds "Population Pressure and Farmland Values in Florida," *Soil and Crop Science Society of Florida Proceedings* Vol 41, 1982, pp 118-21

14 North Central Regional Committee on Land Values *Ongoing Farmland Market Research A Handbook* North Central Regional Research Publication No 306 University of Nebraska, Lincoln, Sept 1985

15 Pasour, E C , Jr "Real Property Taxes and Farm Real Estate Values Incidence and Implications," *American Journal of Agricultural Economics* Vol 55, 1973, pp 549-56

16 _____ "The Capitalization of Real Property Taxes Levied on Farm Real Estate," *American Journal of Agricultural Economics* Vol 57, 1975, pp 539-48

17 Pfeffermann, D , and C H Barnard, "Some New Estimators for Small Area Means with Application to the Assessment of Farmland Values," *Journal of Business and Economic Statistics* (forthcoming) Copies of the manuscript are available from the authors

18 Reynolds, J E "Analytical Studies of Factors Affecting Farm Real Estate Values," *The Farm Real Estate Market Proceedings of a Regional Workshop, May 1984* SNREC Publication No 21, Mississippi State Univ , Apr 1986

19 Reynolds, J E , and D Comer *Southeast Land Value Survey Estimates of Agricultural Land Values for Florida, Alabama, Georgia, and South Carolina, (Final Report)* Cooperative Agreement No 58-319V-5-00375 Univ of Florida, Gainesville, July 1988

20 Reynolds, J E , and D L Tower "Factors Affecting Rural Land Prices in an Urbanizing Area," *Review of Regional Studies* Vol 8, 1978, pp 23-24

21 Reynolds, J E , and R H Tseng "Factors Affecting Farmland Values Among Florida Counties," *Soil and Crop Science Society of Florida* Vol 29 1969, pp 44-48

22 Sandrey, R A L M Arthur, R A Oliveira, and W R Wilson "Determinants of Oregon Farmland Values A Pooled Cross-Sectional, Time Series Analysis," *Western Journal of Agricultural Economics* Vol 7 1982, pp 211-20

23 Scharlach, W C , and G E Schuh "The Land Market as a Link Between the Rural and Urban Sectors of the Economy," *Journal of Farm Economics* Vol 44 1961, pp 1,406-11

24 U S Department of Agriculture, National Agricultural Statistics Service *Agricultural Statistics 1988*, 1988,,pp 377-78

25 U S Department of Agriculture, Economic Research Service and National Agricultural Statistics Service *Farm Costs and Returns Surveys, 1984-1989*

26 U S Department of Agriculture, Economic Research Service *Major Statistical Series of the U S Department of Agriculture Land Values and Land Use* Agriculture Handbook No 671, Vol 6, Aug 1988

27 U S Department of Agriculture, Economic Research Service *Agricultural Land Values and Markets, Outlook and Situation Report* AR-10, June 1988

28 U S Department of Commerce, Bureau of the Census *1982 Census of Agriculture*, 1983

29 Wald, A "Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large," *Transactions of the American Society* Vol 54, 1943, pp 426-82