A SIMULTANEOUS EQUATIONS MODEL OF SPLIT TICKET VOTING IN ISRAELI ELECTIONS[1]

Martin S. Geisel, Carnegie-Mellon University
Naomi E. Kies, Hebrew University
Timothy W. McGuire, Carnegie-Mellon University
Howard Rosenthal, Carnegie-Mellon University

1. Introduction

An election in which citizens vote simultaneously for more than one office is a complex social process where two or more decisions are interrelated. Where more than one office is at stake and where candidates are labeled as to political party voters can vote a "straight" party ticket or they can "split" their ballot, voting for candidates for two or more parties. Politicos have long felt that the two choices were mutually dependent. In this paper we present models of one such election, that of Israel in 1965. Although the choice of the Israeli 1965 data was largely one of circumstance, it was fortunate in that the Israelis voted for only two offices, the national legislature (Knesset) and local municipal councils. We thus avoid the contaminating effects of additional contests. Furthermore, for the city of Jerusalem, much split-ticket voting was suggested by the presence of Teddy Kollek, an attractive candidate at the local level representing the otherwise unsuccessful Rafi party. Our data base consists of voting and census statistics for the sixty-four census tracts of (pre-1967) Jerusalem.

To describe these interrelated phenomena a system of equations is required since a party's municipal vote depends on, among other things, its national vote, and vice versa. The statistical analysis of systems of simultaneous linear equations has received much attention from econometricians. The models with which we shall deal are somewhat more complicated since they involve both non-linearities and inequality constraints on the parameters of the system.

We can introduce our estimation problem through an example involving only a two equation model. Let R_{Ki} denote the proportion of Rafi voters in the i-th district for the Knesset election in 1965; R_{Mi} the corresponding municipal proportion; p's, q's, β's, and γ's constants to be estimated; u's and v's stochastic terms in the model; and let $i = 1,2,\ldots,n$ index the census tracts. Let us first examine a single-equation model originally suggested by Goodman [4] and frequently used in political science applications in the study of two temporally separated elections.[2] For each census tract, assume that the Rafi municipal list maintained a proportion $p_1 + u_{1i}$ of its Knesset vote while winning a proportion $p_2 + u_{2i}$ of the Knesset vote for all other parties (including abstentions). Then

$$\begin{aligned} R_{Mi} &= (p_1+u_{1i})R_{Ki} + (p_2+u_{2i})\ (1-R_{Ki}) \\ &= p_2 + (p_1-p_2)R_{Ki} + u_{2i} + (u_{1i}-u_{2i})R_{Ki}. \end{aligned} \tag{1.1}$$

Let $e_i = u_{2i} + (u_{1i}-u_{2i})R_{Ki}$ and assume the R_{Ki}'s are given numbers. Then, if $E(e_i) = 0$ and $E(e_ie_j) = \sigma^2$ if $i=j$ and zero otherwise, ordinary least squares applied to

$$R_{Mi} = \beta_0 + \beta_1 R_{Ki} + e_i \tag{1.2}$$

yields minimum variance linear unbiased estimates of p_1 and p_2 via the relations $\hat{p}_o = \hat{\beta}_0 + \hat{\beta}_1$, $\hat{p}_2 = \hat{\beta}_0$, where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the least squares estimates from equation (1.2). For the constant variance assumption to hold, either $u_{2i} = u_{1i}$ or $R_{Ki}{}^2\ \mathrm{Var}(u_{1i}) + (1-R_{Ki})^2\ \mathrm{Var}(u_{2i}) + 2R_{Ki}(1-R_{Ki})\ \mathrm{Cov}(u_{1i},u_{2i})$ is constant for all i. Otherwise, procedures which allow for heteroscedastic errors will be required in order to obtain best linear unbiased estimates.

A variant of the Goodman model was suggested by Rosenthal [12]. Instead of letting the proportion won from the remaining parties be a constant, he suggested letting this proportion be a function of the Knesset vote proportion, for example, $p_2/(1-R_{Ki})$. Then,

$$R_{Mi} = (p_1+u_{1i})R_{Ki} + \left(\frac{p_2}{(1-R_{Ki})} + u_{2i}\right)(1-R_{Ki}). \tag{1.3}$$

Simple algebra will show that (1.3) also leads to (1.2) except that the estimates are $\hat{p}_1 = \hat{\beta}_1$, $\hat{p}_2 = \hat{\beta}_0$. The reader can conceive of still further models that could be associated with (1.2). Although the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ may enable one to eliminate some of these models on logical grounds (see Rosenthal [12] and Goodman [4]), any regression poses difficult problems of interpretation.

If we regard R_{Ki} and R_{Mi} both as endogenous (mutually determined by the electoral system) variables, then the least squares estimates of the parameters of (1.2) will be biased and inconsistent since e_i and R_{Ki} will be correlated in general (see, e.g., Johnston [6,pp.231-4]).

To analyze the simultaneous problem another relationship involving R_{Ki} is required. An equation analogous to (1.2) with the roles of R_{Mi} and R_{Ki} interchanged will not do, for the parameters of this two equation system are not estimable or identifiable. The parameters can be identified if we complicate the equations by adding exogenous (determined outside the system) variables to the model. To illustrate this, let

X and Y be two sociological variables (e.g. educational level and proportion of Oriental birth). Modifying (1.1) so as to make the proportions functions of the sociological variables leads, for example, to:

$$R_{Mi} = (p_1+p_3X_i+u_{1i})R_{Ki} + (p_2+p_4X_i+u_{2i})(1-R_{Ki}), \qquad (1.4)$$

$$R_{Ki} = (q_1+q_3Y_i+v_{1i})R_{Mi} + (q_2+q_4Y_i+v_{2i})(1-R_{Ki}). \qquad (1.5)$$

This leads to the structural equations:

$$R_{Mi} = \beta_0 + \beta_1R_{Ki} + \beta_2X_i + \beta_3R_{Ki}X_i + e_i, \quad (1.6)$$

$$R_{Ki} = \gamma_0 + \gamma_1R_{Mi} + \gamma_2Y_i + \gamma_3R_{Mi}Y_i + g_i, \quad (1.7)$$

where the relations between the parameters of (1.4)-(1.5) and (1.6)-(1.7) are immediate.

An interesting feature of (1.6) is that even in this simple case there are variables in the system which are products of endogenous and exogenous variables. This does not typically occur in models considered by econometricians.

In the next section of the paper we give a brief discussion of identification and estimation of systems of linear equations. In section 3 we present methods for dealing with the "adding up" constraint (that is, for each census district the sum of the Municipal proportions is unity, and similarly for the Knesset proportions). In section 4 we discuss the methods employed for dealing with the product variables and parameter constraints. The results of the analysis for the 1965 Jerusalem elections are given in section 5. Concluding remarks are given in section 6.

2. Identification and Estimation of Systems of Linear Equations.[3]

In this section we give a brief exposition of methods of estimation of the parameters of systems of equations of the form:

$$\sum_{\ell=1}^{G} y_{t\ell}\gamma_{\ell g} + \sum_{m=1}^{K} x_{tm}\beta_{mg} = u_{tg}, \qquad t=1,\ldots,T,\ g=1,\ldots,G. \qquad (2.1)$$

The y's are jointly dependent or endogenous variables and the x's are predetermined variables which are stochastically independent of the errors, u_{tg}. The equations (2.1) may be expressed conveniently in matrix notation as

$$Y\Gamma + XB = U, \qquad (2.2)$$

where Y is a $T \times G$ matrix of values of the endogenous variables, X a $T \times K$ matrix of values of predetermined variables, U a $T \times G$ matrix of unobservable stochastic disturbances, and Γ and B are $G \times G$ and $K \times G$ matrices of structural parameters.

We assume that the columns of X are linearly independent so that the rank of X is K.[4] Further, we assume that $|\Gamma| \neq 0$ so we can solve for the reduced form

$$Y = -XB\Gamma^{-1} + U\Gamma^{-1} = X\Pi + V. \qquad (2.3)$$

Regarding the u's we assume

$$E(u_{tg}) = 0, \quad t=1,\ldots,T, g=1,\ldots,G,$$

and

$$E(u_{sg}u_{th}) = \begin{cases}\sigma_{gh} & \text{if } s=t, \\ 0 & \text{if } s\neq t.\end{cases}$$

If we let $u'(t) = [u_{t1},\ldots,u_{tg}]$ we may express the above as

$$E(u'(t)) = [0,\ldots,0],$$

$$E(u(s)u'(t)) = \begin{cases}\Sigma = \{\sigma_{gh}\} & \text{if } s=t, \\ \underset{\sim}{0}_{GG} = \{0\} & \text{if } s\neq t,\end{cases}$$

where Σ is a positive definite symmetric $G \times G$ matrix[5] and $\underset{\sim}{0}_{GG}$ is the null matrix of order G.

The reduced form (2.3) is similar to a multivariate linear regression model. This suggests estimation of the reduced form by ordinary least squares and then estimation of structural parameters using the relation

$$\Pi = -B\Gamma^{-1}.$$

To see that this is impossible in general let us post-multiply the structure (2.2) by an arbitrary non-singular matrix L to give

$$Y\Gamma L + XBL = UL$$

or

$$Y\Gamma^* + XB^* = U^* \qquad (2.4)$$

with reduced form

$$Y = X\Pi^* + V^*,$$

where

$$\Pi^* = -B^*\Gamma^{*-1} = -BL(\Gamma L)^{-1} = -B\Gamma^{-1} = \Pi,$$

so that the two structures have the same reduced form.[6] The two structures (2.2) and (2.4) are observationally equivalent in the sense that the likelihood function for one structure is the same as that for the other. It is thus impossible in general to estimate structural parameters from the reduced form without the imposition of some additional prior information about the structural parameters. This is called the identification problem.[7]

It should be emphasized that identification is not a problem unique to simultaneous equation systems nor is it related to sample size. For example, suppose we are interested in estimating the weights of an apple and an orange. If we always weigh them together we can by repetitive weighing obtain a very precise estimate of the sum of the weights, but regardless of sample size we never can gain any information about the individual weights from this experiment.

We shall discuss only one kind of prior information in relation to the identification problem; namely, knowledge that certain elements of Γ and B are zero. We shall discuss conditions for identification of the parameters of the gth equation.

Let us write the gth equation of the system as

$$y_g = Y_g\gamma_g + X_g\beta_g + u_g, \tag{2.5}$$

where Y_g is a $T \times G_g$ matrix of values of endogenous variables which are included in the gth equation, X_g is a $T \times K_g$ matrix of included exogenous variables, y_g is a $T \times 1$ vector of observations on one endogenous variable which we put on the left hand side, u_g is a $T \times 1$ vector of disturbances, and γ_g and β_g are $G_g \times 1$ and $K_g \times 1$ vectors of parameters, whose elements are the appropriate elements of the gth columns of Γ and B of (2.2), respectively, all divided by $-\gamma_{gg}$.

Now write the part of the reduced form corresponding to y_g and Y_g as

$$[y_g:Y_g] = [X_g:X_g^*] \begin{bmatrix} \pi_g & \Pi_g \\ \pi_g^* & \Pi_g^* \end{bmatrix} + \text{reduced form disturbances},$$

where X_g^* is the $T \times (K-K_g)$ matrix of exogenous variables excluded from (2.5), π_g is $K_g \times 1$, π_g^* is $(K-K_g) \times 1$, Π_g is $K_g \times G_g$, and Π_g^* is $(K-K_g) \times G_g$. If we post-multiply by $[1: -\gamma_g']'$ we have on the left hand side

$$y_g - Y_g\gamma_g = X_g\beta_g + u_g ,$$

where the equality is by (2.5). Now, neglecting disturbances,[8] we must have on the right

$$[\pi_g:\Pi_g]\ [1: -\gamma_g']' = \beta_g, \tag{2.6}$$

$$[\pi_g^*:\Pi_g^*]\ [1: -\gamma_g']' = 0; \tag{2.7}$$

(2.6) gives β_g in terms of γ_g given π_g and Π_g. It does not imply restrictions on π_g and Π_g. Relation (2.7) says

$$\pi_g^* = \Pi_g^*\gamma_g .$$

This system of equations has a unique solution for γ_g if and only if the ranks of Π_g^* and $[\pi_g^*: \Pi_g^*]$ are both G_g. This is the <u>rank condition</u> for identification. This rank can be G_g only if $K - K_g \geq G_g$; that is, only if the number of exogenous variables excluded from the gth relation is greater than or equal to the number of endogenous variables in the equation minus one. This necessary condition is called the <u>order condition</u>. Both conditions may easily be extended to the case of homogeneous linear restrictions on the parameters of the gth equation. The rank condition may also be stated in terms of the structural coefficient matrices (see Johnston [6, p.251]).

Having established criteria for the possibility of estimation, we now proceed to discuss estimation methods. If ordinary least squares were applied directly to (2.5) the estimators would be inconsistent because of the correlation between Y_g and u_g. Theil [14,15] developed the following approach which yields consistent estimators and is known as <u>two stage least squares</u> (2SLS).[9]

Now Y_g is a submatrix of $Y = -XB\Gamma^{-1} + U\Gamma^{-1}$, so in (2.5) let us replace Y_g with its equivalent expression from (2.3); that is,

$$y_g = [(-XB\Gamma^{-1})_g : X_g] \begin{bmatrix} \gamma_g \\ \beta_g \end{bmatrix} + u_g + (U\Gamma^{-1})_g\gamma_g .$$

However, $B\Gamma^{-1}$ is unknown. Nevertheless, we may estimate Y_g by ordinary least squares from the reduced form:

$$\hat{Y}_g = X(X'X)^{-1}X'Y_g = X\hat{\Pi}_g .$$

Then

$$Y_g = X\hat{\Pi}_g + V_g ,$$

where

$$V_g = [I - X(X'X)^{-1}X']Y_g .$$

This is the first stage. The second stage consists of regressing y_g on $X\hat{\Pi}_g$ and X_g; that is,

$$y_g = [Y_g - V_g : X_g]\ [\gamma_g' : \beta_g']' + u_g + V_g\gamma_g .$$

Then

$$(\hat{\gamma}_g' : \hat{\beta}_g')' = M_{\hat{z}\hat{z}}^{-1}\, m_{\hat{z}y_g} ,$$

where

$$M_{\hat{z}\hat{z}} = \begin{bmatrix} (Y_g-V_g)'(Y_g-V_g) & (Y_g-V_g)'X_g \\ X_g'(Y_g-V_g) & X_g'X_g \end{bmatrix}$$

and

$$m_{\hat{z}y_g} = [(Y_g - V_g):X_g]'y_g \ .$$

For the inverse of the matrix $M_{\hat{z}\hat{z}}$ to exist it is necessary that the order condition be satisfied. Note that $X_g'V_g = \underset{\sim}{0}_{K_g G_g}$ (the $K_g \times G_g$ null matrix) and $(Y_g-V_g)'(Y_g-V_g) = Y_g'Y_g - V_g'V_g$. The The 2SLS estimator has a normal limiting distribution with mean $[\gamma_g':\beta_g']'$ and covariance matrix $\sigma_{gg}M_{\hat{z}\hat{z}}^{-1}$ [however, finite sample moments need not exist (see, e.g., Dhrymes [2,pp. 180, 204])]; σ_{gg} may be estimated consistently by

$$s_{gg} = \frac{1}{T}(y_g - Y_g\hat{\gamma}_g - X_g\hat{\beta}_g)'(y_g - Y_g\hat{\gamma}_g - X_g\hat{\beta}_g).$$

The method of <u>three stage least squares</u> (3SLS) is a method for simultaneous estimation of the entire system. If we let $Z_g = [Y_g:X_g]$ and $\delta_g = [\gamma_g':\beta_g']'$ and premultiply each equation by X' we may write

$$\begin{bmatrix} X'y_1 \\ \cdot \\ \cdot \\ \cdot \\ X'y_G \end{bmatrix} = \begin{bmatrix} X'Z_1 & 0 \ldots 0 \\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot & \cdot\ \cdot\ \cdot \\ \cdot & \cdot\ \cdot\ \cdot \\ 0 & 0 \ldots X'Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \cdot \\ \cdot \\ \cdot \\ \delta_G \end{bmatrix} + \begin{bmatrix} X'u_1 \\ \cdot \\ \cdot \\ \cdot \\ X'u_G \end{bmatrix}$$

or

$$w = W\delta + \epsilon \ ,$$

with

$$E\,\epsilon\,\epsilon' = \Sigma \otimes X'X.$$

Then the 3SLS estimator of δ is

$$\hat{\delta} = \{W'[\Sigma^{-1} \otimes (X'X)^{-1}]W\}^{-1}\, W'[\Sigma^{-1} \otimes (X'X)^{-1}]w, \quad (2.9)$$

with the unknown matrix Σ replaced by an estimate of itself constructed from the 2SLS residuals. The estimator $\hat{\delta}$ is consistent, has covariance matrix approximately equal to $\{W'[\Sigma^{-1} \otimes (X'X)^{-1}]W\}^{-1}$, and is asymptotically efficient unless some of the elements of Σ are known (see Rothenberg and Leenders [13] and Madansky [10]).

3. The Adding-Up Conditions

<u>Implications for the reduced form</u>. Denote the typical reduced form observation by

$$\underset{(1\times G)}{y'(t)} = \underset{(1\times K)}{x'(t)}\ \underset{(K\times G)}{\Pi} + \underset{(1\times G)}{v'(t)}, \quad t=1,\ldots,T, \quad (3.1)$$

with

$$y'(t)\, \underset{\sim}{1}_G = 1 \quad \left(\sum_{i=1}^{G} y_i(t) = 1\right), \quad t=1,\ldots,T, \quad (3.2)$$

where $\underset{\sim}{1}_G$ is a G-component (column) vector with each element equal to unity. Clearly

$$y'(t)\,\underset{\sim}{1}_G = x'(t)\Pi\underset{\sim}{1}_G + v'(t)\underset{\sim}{1}_G\ , \quad t=1,\ldots,T,$$

so from (3.2),

$$1 = x'(t)\Pi\underset{\sim}{1}_G + v'(t)\underset{\sim}{1}_G, \quad t=1,\ldots,T; \quad (3.3)$$

taking expectations,

$$x'(t)\Pi\underset{\sim}{1}_G = 1, \quad t=1,\ldots,T. \quad (3.4)$$

Substituting this result back into (3.3),

$$v'(t)\underset{\sim}{1}_G = 0, \quad t=1,\ldots,T; \quad (3.5)$$

thus

$$\Omega\, \underset{\sim}{1}_G = \underset{\sim}{0}_G\ ,$$

where

$$\Omega = E[v(t)\ v'(t)], \quad t=1,\ldots,T.$$

Let us examine the restrictions on Π more closely. First, it follows immediately from (3.2) and the fact that

$$\Pi_{\cdot} \equiv \Pi\ \underset{\sim}{1}_G = (\pi_{1\cdot}, \pi_{2\cdot}, \ldots, \pi_{K\cdot})'$$

(where $\pi_{i\cdot} = \sum_j \pi_{ij}$) is a constant vector that there exists at least one linear combination of the columns of X that sums to a given non-zero number (say unity), and the condition that X be of full column rank (K) implies that there can be at most one such linear combination; consequently there is exactly one such combination, and hereinafter it is assumed that $x_{t1} = 1$, $t=1,\ldots,T$.

Next it is shown that $\pi_{i\cdot} = 0$, $i=2,\ldots,K$. Condition (3.4) may be written equivalently as

$$\sum_{i=1}^{G} x_{ti}\pi_{i\cdot} = 1, \quad t=1,\ldots,T.$$

Suppose this restriction holds for a given X; then it cannot hold if we perturb any x_{ti} for which $\pi_{i\cdot} \neq 0$. Since the x_{ti}'s are unconstrained for $i=2,\ldots,G$, it follows that $\pi_{i\cdot} = 0$, $i=2,\ldots,G$. Finally, it follows that $\pi_{1\cdot} = 1$. Although stated differently, these conditions are essentially equivalent to those stated in McGuire <u>et al</u>. [11].

It should be noted that if $y'(t)$ is subject to inequality constraints of the form

$$\underset{\sim}{0}'_G \le y'(t) \le \underset{\sim}{1}'_G, \quad t=1,\ldots,T,$$

then

$$\underset{\sim}{0}'_G \le x'(t)\ \Pi \le \underset{\sim}{1}'_G\ , \quad t=1,\ldots,T. \quad (3.6)$$

One obvious implication of this condition is that the $x'(t)$'s must be bounded, for if one or more components of $x'(t)$ is unbounded, then it is always possible to choose a value great enough that (3.6) is violated.

Implications for the structural equations. Define

$$\Gamma^{\cdot} = \Gamma^{-1}\underset{\sim}{1}_G = (\gamma^{1\cdot}, \gamma^{2\cdot}, \ldots, \gamma^{G\cdot})' ,$$

where $\gamma^{i} = \sum_j \gamma^{ij}$ is the i,j-th element of Γ^{-1}.

It was shown above that $[1:\underset{\sim}{0}'_{K-1}]' = \Pi\, \underset{\sim}{1}_G$; in terms of restrictions on the coefficient matrix B in the structural equations this condition becomes

$$[1:\underset{\sim}{0}'_{K-1}]' = \Pi\underset{\sim}{1}_G = (\Pi\Gamma)(\Gamma^{-1}\underset{\sim}{1}_G) = -B\Gamma^{\cdot}.$$

Similarly, (3.5) implies

$$0 = v'(t)\underset{\sim}{1}_G = (v'(t)\Gamma)(\Gamma^{-1}\underset{\sim}{1}_G)$$
$$= u'(t)\Gamma^{\cdot} , \qquad t=1,\ldots,T;$$

thus

$$\underset{\sim}{0}_G = \Omega\, \underset{\sim}{1}_G = (\Gamma')^{-1}\Gamma'\Omega\Gamma\Gamma^{-1}\underset{\sim}{1}_G$$
$$= (\Gamma')^{-1}\Sigma\, \Gamma^{\cdot} = \Sigma\, \Gamma^{\cdot};$$

hence, the rank of Σ, which is equal to the rank of Ω since Γ has full rank (G), is less than G.

The linear dependence of the elements of the disturbance vector impinges on the identification of the elements of Γ. Define

$$z^{(i)}(t) = (z_{t,1}, \ldots, z_{t,i-1}, z_{t,i+1}, \ldots, z_{t,m})',$$

where $z(t)$ is an m-component (column) vector, and partition the structural model

$$y'(t)\Gamma + x'(t)B = u'(t), \qquad t=1,\ldots,T,$$

$$(y^{(G)'} : y_G) \begin{bmatrix} \Gamma_{11} & \Gamma_{21} \\ & \\ \Gamma_{12} & \Gamma_{22} \end{bmatrix} + (B_{11} : B_{21})$$
$$+ x^{(1)'}[B_{12} : B_{22}]$$
$$= (u^{(G)'} : u_G),$$

where the nature of the partitioning is obvious. Using (3.2), this relationship may be written equivalently as

$$y^{(G)'}[\Gamma_{11} - \underset{\sim}{1}\Gamma_{12} : \Gamma_{21} - \underset{\sim}{1}\Gamma_{22}]$$
$$+ (B_{11} + \Gamma_{12} : B_{21} + \Gamma_{22})$$
$$+ x^{(1)'}[B_{12} : B_{22}] = (u^{(G)'} : u_G).$$

Define

$$\bar{\Gamma}_{11} = \Gamma_{11} - \underset{\sim}{1}\Gamma_{12}, \quad \bar{\Gamma}_{21} = \Gamma_{21} - \underset{\sim}{1}\Gamma_{22} ,$$

$$\bar{B}_{11} = B_{11} + \Gamma_{12}, \quad \bar{B}_{21} = B_{21} + \Gamma_{22};$$

then this system is described completely by

$$y^{(G)}\bar{\Gamma}_{11} + \bar{B}_{11} + x^{(1)}B_{12} = u^{(G)} \qquad (3.7)$$

since $\bar{\Gamma}_{21}$ is a linear combination of the columns of $\bar{\Gamma}_{11}$. This set of structural relations is in the form of standard classical econometric models and can be analyzed as such (i.e., the identification conditions and estimation procedures described in the previous section apply directly to this specification). It should be obvious that identification of $\bar{\Gamma}_{11}$, $\bar{B}_{11}$ and B_{12} is the most that possibly can be achieved in this model, for model (3.7) assumes that form for all Γ_{12}, Γ_{21}, and Γ_{22}.

4. Non-Linear Estimation and Identification

In the previous sections we treated identification, estimation, and the adding-up condition in linear simultaneous equations systems. We now briefly discuss these areas with respect to the non-linear problem to be treated here.

The problem of identification of certain non-linear systems has been treated by Fisher [3] and Kelejian [7]. Consider a system of simultaneous equations

$$y'(t)\Gamma + F'(t)A + x'(t)\, B = u'(t),$$

where $F'(t)$ is a vector of non-linear functions of contemporaneous endogenous and exogenous variables, and let $H(t) = E[F(t)]$. Kelejian proves the following result. If the columns of X and $F = [F(1), \ldots, F(T)]'$ are linearly independent, "each additional endogenous function may be considered, for identification purposes, as just another linearly independent predetermined variable" [Kelejian, 7, p.7]. By this rule, all the equations in non-linear systems we shall consider in this study are over-identified.

Estimation poses more of a problem. The technique used here is to approximate the reduced form with a second order Taylor Series expansion, giving the i-th reduced form equation[10]

$$y_{ti} = \sum_{k=1}^{K}\sum_{j=1}^{K} x_{tk}x_{tj}\beta_{ij} + v_{ti}, \qquad \begin{matrix} i=1,\ldots,G, \\ t=1,\ldots,T. \end{matrix} \qquad (4.1)$$

The predicted values of the endogenous variables from (4.1) are then used in the second and third stages.

There are two sets of adding-up conditions for each observation (census tract) in our problem: the proportions of the Knesset vote received by each party (including abstentions) sum to unity, as do the corresponding Municipal proportions. Accordingly, we eliminate one Knesset equation and one Municipal equation prior to estimation.

5. Results for the 1965 Jerusalem Elections

In this section we analyze a model based on a ten equation model. The endogenous variables in these structural relations are the proportions of the registered voters casting ballots for the (1) Rafi, (2) Gahal, (3) All Religious, (4) All Other Secular [about three-fourths of which consists of the Alignment (of Mapai and Ah'dut A'avoda) vote] parties and (5) abstaining in the Knesset and Municipal races. Since this model is primarily illustrative and is not what we regard as the best specification of the election, we aggregated the All Religious, All Other Secular, and Abstention votes into an All Other category to prevent the analysis from becoming too cumbersome. This aggregation is not likely to conceal any interesting switching among religious parties, since Kies and Rosenthal [9] have shown that the squared correlations between the Knesset and Municipal votes received by the three religious parties, Poalei Aguda, NRP, and Aguda, are 0.92, 0.95, and 0.98, respectively; similar results were obtained for abstentions. If there are any explanatory variables influencing the All Religious, All Other Secular, or Abstention votes which do not drop out when these variables are aggregated into the All Other category and which are not included in at least the Rafi or Gahal pair of structural relations, then the absence of such variables is a specification error.

We chose to retain the pair of Rafi equations for this illustrative example because we hypothesized that the feedback effects were greatest for this party due to (1) Kollek's presence and (2) Rafi being a new party. Gahal was retained because it was shown in Kies and Rosenthal [9] that, at least with OLS estimation, the Gahal municipal equation benefitted substantially from the inclusion of "product" variables as against ordinary linear variables.

In any simultaneous equations problem, the choice of exogenous variables is obviously limited by available data. We had the results of the 1961 elections, when only the Knesset was elected, and detailed census data (occupational structure, ethnic origin broken down by period of immigration, sex, age, housing conditions, etc.). Unfortunately, we did not have data for religious practice, one of the sociological variables which may play an important role in political analysis.

It seems important to use the 1961 Knesset vote variables since voter behavior will reflect organizational and historical as well as socioeconomic influences on preferences. In fact, analysis presented in Kies [8] indicates that in this case past vote variables are generally better predictors than socioeconomic variables and that socioeconomic variables frequently add little to explained variance over what is "explained" by the past vote variables.

In fact, we used only two socioeconomic variables in the analysis chosen as markers for the two major dimensions found in a Guttman-Lingoes SSA-I analysis for the socioeconomic variables (see Kies [8]). The two variables were the proportion of the total population immigrant from "Oriental" countries (chiefly Sephardic Jews from North Africa and the Middle East) before 1948, which we denote by Seph. 48-,and the proportion of the adult population with less than one complete year of formal schooling, denoted by Low Ed.; both variables come from the 1961 census.

Since the original ten-equation model has been reduced to a six-equation model by aggregation and since one Knesset equation and Municipal equation are estimated indirectly by using the "adding up" conditions, we estimated four equations. We chose to eliminate the pair of All Other relations, so the structural equations we estimated are for the (1) Rafi Knesset, (2) Rafi Municipal, (3) Gahal Knesset, and (4) Gahal Municipal proportions, respectively. The OLS and 3SLS estimates of the model parameters are presented in Table 1. We also estimated a linearized version of the model; the 3SLS estimates of the parameters and the implied reduced form (RF) paramter estimates (i.e., $\hat{\Pi} = \hat{B}\hat{\Gamma}^{-1}$ in the notation of sections 2 and 3) are presented in Table 2.

Comparison of the estimation methods. The OLS estimates would seem to be reasonably satisfactory insofar as the R^2 values, which range from .80 to .93 for the nonlinear model (.76 to .92 for the linear model, although these estimates are not shown due to space limitations) are evidence of good cross-sectional specification. However, since the equations contain endogenous variables as regressors, the estimates are inconsistent. The 3SLS estimates, which are consistent, also indicate good explanatory power. The importance of using 3SLS is evidenced by the differences in the estimates obtained using the two procedures, which in most cases exceed one standard error (using either the OLS or the 3SLS estimate of the standard error of the parameter estimates).

The reduced form estimates. The reduced form estimates for the linear model are consistent with what we know about the structures of the parties and the relationships between the 1961 and 1965 parties. The Rafi party was formed by Mapai dissidents; in the 1965 elections Rafi captured about 55 percent of the Mapai 1961 vote in the Municipal election and 25 percent in the Knesset election. Gahal, a coalition of the Herut and part of the Liberal organizations, picked up 42 percent of the Herut 1961 vote in the Municipal election and 71 percent in the Knesset election while retaining 25 percent of the Liberal 1961 vote in the Municipal election and 42 percent in the Knesset election. It is

TABLE 1. ORDINARY AND THREE STAGE LEAST SQUARES ESTIMATES OF THE NON-LINEAR MODEL

	Equation							
	Rafi-Knesset		Rafi-Municipal		Gahal-Knesset		Gahal-Municipal	
Regression Variable	OLS	3SLS	OLS	3SLS	OLS	3SLS	OLS	3SLS
Constant	0.0114 (0.0049)	0.0072 (0.0053)	0.0053 (0.0115)	-0.0014 (0.0112)	0.0352 (0.0112)	0.0417 (0.0103)	0.0053 (0.0088)	-0.0043 (0.0097)
Herut, 1961					0.3054 (0.0581)	0.2126 (0.0579)		
Liberal, 1961	-0.1296 (0.0429)	-0.1972 (0.0443)	0.4051 (0.0702)	0.4443 (0.0694)	0.1947 (0.0650)	0.1091 (0.0584)		
Mapai, 1961			0.2200 (0.0604)	0.2651 (0.0639)				
Low Ed.					-0.0586 (0.0277)	-0.0920 (0.0292)		
Seph. 48-	-0.1650 (0.0427)	-0.1267 (0.0381)			-0.0721 (0.1053)	-0.2102 (0.1172)		
Rafi-Knesset			1.3281 (0.1500)	1.1179 (0.1936)				
Rafi-Municipal	0.2625 (0.0440)	0.3600 (0.0495)						
Gahal-Knesset							0.4618 (0.0524)	0.5245 (0.0578)
Gahal-Municipal					0.9429 (0.1058)	1.1438 (0.1312)		
Herut, 1961 x Rafi-M	1.0499 (0.1935)	0.7685 (0.1817)						
Low Ed. x Gahal-K			-0.2534 (0.0904)	-0.1006 (0.0827)			0.4286 (0.0559)	0.3757 (0.0566)
Seph. 48- x Gahal-K							1.1975 (0.2172)	1.0381 (0.2203)
Adjusted variance x 10000[a]	1.3881	1.4222	5.7244	6.5218	3.6629	3.7132	2.1404	2.1412
Var. of dep. var. x 10000	6.5825	6.5825	35.760	35.760	38.507	38.507	30.351	30.351

Note a. The adjusted variances for the 3SLS estimates are based on the 2SLS estimates.

interesting to note that Kollek's strategy for cultivating the vote of Sephardic Jews (he gave *Sefaradim* two of the first five places on his list) apparently failed; he captured none of the Herut 1961 vote and, *ceteris paribus*, lost about three-eighths of a vote for every voting Sephardic Jew. (There appears to be little difference between the attractiveness of the Rafi Municipal and Knesset lists to Sephardic Jews; the strong ethnic appeal of Gahal is evidenced by the large positive coefficients of Seph. 48- in the Gahal equations). However, his appeal to high status voters obviously paid off; he captured 36 percent of the Liberal 1961 vote while this variable actually exerted a slight negative effect on the Rafi Knesset vote. Although the educational achievement of the constituency appears not to have been terribly important, its effect is in the predicted direction. It often is argued that independents are more likely to appeal to better educated, "thinking" voters [9]. The negative coefficients of "Low Ed." (proportion of voters completing less than one year of formal education), with the more negative occurring in the Rafi Municipal equation, and the positive coefficients in the Gahal equations are consistent with this hypothesis.

Estimates of a proportions model. In this section we examine the implications of the parameter estimates in the context of a proportions model such as that discussed in the Introduction.

TABLE 2. REDUCED FORM AND THREE STAGE LEAST SQUARES ESTIMATES OF THE LINEAR MODEL

	Equation							
	Rafi-Knesset		Rafi-Municipal		Gahal-Knesset		Gahal-Municipal	
Regression Variable	RF	3SLS	RF	3SLS	RF	3SLS	RF	3SLS
Constant	-0.0033	-0.0033 (0.0077)	0.0001	0.0045 (0.0162)	0.0324	0.0384 (0.0103)	-0.0049	-0.0243 (0.0085)
Herut, 1961	0.0089	0.0095 (0.0286)	-0.0014		0.7102	0.1947 (0.0559)	0.4244	
Liberal, 1961	-0.0585	-0.2251 (0.0469)	0.3606	0.4374 (0.0701)	0.4231	0.1160 (0.0558)	0.2529	
Mapai, 1961	0.2532		0.5481	0.2464 (0.0631)				
Low Ed.	-0.0386		-0.0835	-0.0365 (0.0254)	0.0613	-0.1010 (0.0292)	0.1336	0.0970 (0.0151)
Seph. 48-	-0.3094	-0.1350 (0.0445)	-0.3775		0.5254	-0.2652 (0.1176)	0.6511	0.3370 (0.0710)
Rafi-Knesset				1.1915 (0.1935)				
Rafi-Municipal		0.4620 (0.0433)						
Gahal-Knesset				-0.0168 (0.0623)				0.5977 (0.0531)
Gahal-Municipal						1.2144 (0.1282)		
Adjusted variance x 10000 (see Table 1, Note a)		1.6559		6.1791		3.7132		2.3374
Variance of dependent variable x 10000	6.5825	6.5825	35.760	35.760	38.507	38.507	30.351	30.351

In constructing proportions interpretations we assume: (a) all voters who voted for the Rafi Knesset list also voted for the Rafi Municipal list; (b) all voters who voted for the Gahal Municipal list also voted for the Gahal Knesset list. These severe assumptions obviously are simplifications, allowing for switching in only one direction for each party. Nevertheless, they may not be grossly incorrect in the context of this election, and the parameter estimates of the implied proportions model do not violate any theoretical constraints (the terms in brackets are the proportions).

$$\begin{aligned} \text{Rafi}_K = [&.3600 + .7685(\text{Herut}_{61}) \\ &+ \{.0724 - .1972(\text{Lib}_{61}) \\ &- .1267(\text{Seph. 48-})\}/\text{Rafi}_M]\ \text{Rafi}_M \end{aligned}$$

$$\begin{aligned} \text{Gahal}_K = [&1.0]\ \text{Gahal}_M \\ &+ [\{.1855 + .2126(\text{Herut}_{61}) \\ &+ .1091(\text{Lib}_{61}) - .2102(\text{Seph. 48-}) \\ &- .0920(\text{Low Ed.}) \\ &- .1438(\text{Gahal}_M)\}/(\text{Other}_M + \text{Rafi}_M)] \\ &\times (\text{Other}_M + \text{Rafi}_M) \end{aligned}$$

$$\begin{aligned} \text{Rafi}_M = [&1.0]\ \text{Rafi}_K \\ &+ [\{.2651(\text{Mapai}_{61}) + .4443(\text{Lib}_{61}) \\ &+ .1179(\text{Rafi}_K) \\ &- .1006(\text{Low Ed.} \times \text{Gahal}_K)\}/ \\ &(\text{Other}_K + \text{Gahal}_K)](\text{Other}_K + \text{Gahal}_K) \end{aligned}$$

$$\begin{aligned} \text{Gahal}_M = [&.0043/(\text{Gahal}_K) + .5245 + .3757(\text{Low Ed.}) \\ &+ 1.0381(\text{Seph. 48-})]\ \text{Gahal}_K \end{aligned}$$

A couple of remarks are in order.

(i) Although we need to undertake an

investigation as to whether the proportions are bounded between zero and one for all sample values, investigation of these expressions for various trial values suggests that the functions for the proportions behave appropriately.

(ii) That the various proportions are the complex expressions in brackets reflects the finding that simple proportions models lead to estimates far outside the [0, 1] interval (see Kies and Rosenthal [9]). That we reject the simple proportions model is hardly a reason for accepting the present estimates. Given that the present model is intended as an illustrative example, great caution should be placed on the use of these proportions models. On the other hand, this complexity in "proportions" interpretations of the results would seem to lend further support to the use of simultaneous models.

6. Directions for Further Research

One of the present difficulties of econometric analyses of split ticket voting is the absence of well-developed mathematical theory of voting behavior in multiple-office elections. Progress in empirical analysis is not likely to be made without prior improvements in theory. Accordingly, we now outline the directions which we believe theoretical advances might take and the difficulties likely to be encountered. First we discuss an alternative specification of coattails. We then discuss the problems of modeling cognitive dissonance. Finally, we consider the relationship between the coattails and the cognitive dissonance specifications.

Coattails revisited. For expositional clarity let us consider the pair of equations for Rafi only, which in the absence of coattails may be written

$$R_K = x_1\beta_1 + x_2\beta_2 + \epsilon_K, \tag{6.1}$$

$$R_M = x_1\alpha_1 + x_3\alpha_3 + \epsilon_M. \tag{6.2}$$

Noticeably absent from these equations is a variable measuring the special appeal of a Teddy Kollek. Thus, if we knew the true values of the vectors α_1 and α_2 (or had independent estimates of them) we surely would underpredict R_M. A reasonable estimate of the personal appeal of Kollek is the difference between the vote he received and the vote he was expected to receive; this difference is simply ϵ_M [or, equivalently, $(R_M - x_1\alpha_1 - x_3\alpha_3)$]. One way to model Kollek's coattails is to add the term $\beta_4\epsilon_M$ to (6.1), giving

$$R_K = x_1\beta_1 + x_2\beta_2 + \beta_4\epsilon_M + \epsilon_K. \tag{6.3}$$

We currently are attempting to gather relevant data for one or more other Israeli cities so that we might obtain independent estimates of the α's which can be used in estimating (6.3).

Cognitive dissonance. Consider a two-party [say Rafi (1) and Gahal (2)], two-office [say Knesset (1) and Municipal (2)] election. One way to model the simultaneous choice problem is to assume that each individual first decides how he would vote for each office if that were the only decision and then evaluates the set of independent choices later to check for cognitive balance. Let p^*_{ij} be the proportion of individuals who would vote for party i in election 1 and party j in election 2 if these decisions were treated as independent and let p_{ij} be the corresponding proportion in the actual simultaneous choice situation; also, let R^*_K and R^*_M be the vote proportions which would have been obtained in the hypothetical independent choice situation. Then

$$R^*_K = p^*_{11} + p^*_{12}\ , \quad R^*_M = p^*_{11} + p^*_{21}\ ,$$

$$R_K = p_{11} + p_{12}\ , \quad R_M = p_{11} + p_{21}\ .$$

Now assume that because of political realties, party loyalty, or cognitive dissonance (these concepts are not identifiable at this level of analysis), individuals voting for different parties in the two elections (that is, those voters comprising the p^*_{12} and p^*_{21} terms) reconsider their decisions. It seems reasonable to assume

$$p_{11} = p^*_{11} + \alpha_1 p^*_{12} + \alpha_2 p^*_{21}\ , \quad p_{12} = (1-\alpha_1-\beta_1)p^*_{12},$$

$$p_{22} = p^*_{22} + \beta_1 p^*_{12} + \beta_2 p^*_{21}\ , \quad p_{21} = (1-\alpha_2-\beta_2)p^*_{21},$$

$$0 \le \alpha_1,\alpha_2,\beta_1,\beta_2 \le 1, \quad \alpha_1+\beta_1 \le 1, \quad \alpha_2+\beta_2 \le 1\ .$$

Then it is easily shown that

$$R_K - R^*_K = \alpha_2 p^*_{21} - \beta_1 p^*_{12} \gtreqless 0\ ,$$

$$R_M - R^*_M = \alpha_1 p^*_{12} - \beta_2 p^*_{21} \gtreqless 0\ .$$

Thus at this level theory provides no clues about the magnitude or direction of adjustments due to cognitive dissonance; indeed, it even provides no clues about the variables that should be in the model.

When we move on to two elections each involving more than two parties the difficulties are compounded. For example, consider the individual whose independent choices are Gahal in the Knesset election and Independent Liberal in the Municipal election. One reasonable way (in addition to the previously discussed ways) that the individual might achieve cognitive balance is to vote for Rafi in both elections.

The prospects for modeling cognitive dissonance adjustments in an aggregate model are dim. Apparently the first two sentences in chapter 11 of Brown[11] should not be overlooked: "The general experimental design for discovering determinants of attitude change is a simple one. Some sort of an attitude must be measured before and after [italics ours] the interpolation of persuasive communications which differ from one another in some known respect."

Coattails and cognitive dissonance. It would appear that the coattails phenomenon is a subset of cognitive dissonance. What we mean by an unusually attractive candidate is one who can attract votes which his party ordinarily would not have received. His ability to carry some portion of these votes for the party in elections for other offices would seem to result from the cognitive balancing on the part of the extra individuals who were attracted by the man (as opposed to the party). The practical difference between coattails in particular and cognitive dissonance in general is that in the former case we believe that we have a method for dealing with the situation empirically at an aggregate level while in the latter case we do not.

FOOTNOTES

[1]This research was supported in part by National Science Foundation grant GS-2945 to Howard Rosenthal and NSF grant GS-2751 to Timothy McGuire (and others).

[2]We emphasize that this approach abstracts from the simultaneity problem which is of central interest here. We include a discussion of it for the sake of continuity with earlier literature and to motivate the simultaneous problem.

[3]This subject is discussed in detail in any econometrics text. A good introductory exposition is given in Johnston [6].

[4]We also assume $\operatorname*{plim}_{T \to \infty} \frac{1}{T} X'X = Q$ with $|Q| \neq 0$. This assumption guarantees that the estimation procedures discussed below have desirable asymptotic properties.

[5]In many econometric models, some equations are identities with no disturbance terms. In this case Σ is non-negative definite.

[6]A simple counting argument will also indicate this. Π contains KG elements whereas the total number of elements in B and Γ is $KG + G^2$.

[7]For a general treatment of identification see Fisher [3]. For a Bayesian approach see Zellner [16].

[8]It is straightforward to show that this function of the reduced form disturbances is exactly u_g.

[9]The method was independently developed by Basmann [1]. Also, a substantial amount of work has been done on maximum likelihood methods (see Hood & Koopmans [5]). For a more recent exposition of estimation methods see Dhrymes [2].

[10]In actual fact we omitted three variables, (Seph. 48-x Liberal, 1961), (Low Ed. x Mapai, 1961), and (Herut, 1961 x Liberal, 1961), due to the high degree of collinearity in the set of first and second order exogenous variables. Since the number of such variables is (K+1)K/2, the multi-collinearity problem is likely to be serious in this situation. One possibility might be to use principal components analysis to eliminate unimportant variables.

[11]Roger Brown, Social Psychology (New York: The Free Press, 1965), p. 549.

REFERENCES

[1] Basmann, R. L., "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," Econometrica, 25 (1957), 77-83.

[2] Dhrymes, P. J., Econometrics. New York: Harper & Row, 1970.

[3] Fisher, F. M., The Identification Problem in Econometrics. New York: McGraw-Hill, 1966.

[4] Goodman, L. A., "Some Alternatives to Ecological Correlation," American Journal of Sociology, 64 (1959), 610-25.

[5] Hood, W. C. and Koopmans, T. C., eds., Studies in Econometric Method. New York: John Wiley & Sons, Inc., 1953.

[6] Johnston, J., Econometric Methods. New York: McGraw-Hill Book Company, Inc., 1963.

[7] Kelejian, H. H., "Identification of Nonlinear Systems: An Interpretation of Fisher (also, the Multiple Solution Problem)," presented at the Second World Congress of the Econometric Society, Cambridge, England, September, 1970.

[8] Kies, N. E., Constituency Support and the Israeli Party System: An Analysis of Elections in Jerusalem, 1959-65. Unpublished Ph.D. dissertation, MIT, 1969.

[9] Kies, N. E. and Rosenthal, H., "Split Tickets and Aggregate Analysis: Jerusalem, 1965, unpublished Carnegie-Mellon University Working Paper WP-28-69-7, 1969.

[10] Madansky, A., "On the Efficiency of Three-Stage Least-Squares Estimation," Econometrica, 32 (1964), 51-56.

[11] McGuire, T. W., Farley, J. U., Lucas, R. E. and Ring, L. W., "Estimation and Inference for Linear Models in Which Subsets of the Dependent Variable are Constrained," Journal of the American Statistical Association, 63 (1968), 1201-13.

[12] Rosenthal, H., "Aggregate Data," in I. Pool et al., eds., Handbook of Communications, Chicago: Rand McNally, forthcoming

[13] Rothenberg, T. J. and Leenders, C. T., "Efficient Estimation of Simultaneous Equation Systems," Econometrica, 32 (1964), 57-76.

[14] Theil, H., Economic Forecasts and Policy, Second Revised Edition. Amsterdam: North-Holland Publishing Company, 1961.

[15] ________, "Estimation and Simultaneous Correlation in Complete Equation Systems," The Hague: Central Planning Bureau, 1953.

[16] Zellner, A., "Bayesian Inference in Econometrics," unpublished University of Chicago manuscript, 1970.