EFFICIENT ESTIMATION OF THE GEOMETRIC DISTRIBUTED LAG MODEL; SOME MONTE CARLO RESULTS ON SMALL SAMPLE PROPERTIES

F.C. Palm
E. Vogelvang
D.A. Kodde

Researchmemorandum 1980-12

Amsterdam, November 1980

Efficient Estimation of the Geometric Distributed Lag Model
Some Monte Carlo Results on Small Sample Properties.

F.C. Palm
E. Vogelvang
D.A. Kodde

Abstract

In this paper we report Monte Carlo results on the small sample properties of instrumental variables, asymptotically efficient two-step and iterative Gauss-Newton estimators for a Koyck (1954) distributed lag model with uncorrelated errors (model 1) and with first order autoregressive errors (model 2). We use the technique of control variables to increase the precision of the Monte Carlo results and summarize the outcome using response functions.

Two main questions have been investigated for a sample size T=30 and T=60:
(a) are the asymptotically efficient estimators to be preferred to a consistent but inefficient instrumental variables estimator?,
(b) does it pay to iterate an asymptotically efficient estimator until convergence is achieved?

For the sample sizes considered, we conclude that the efficient two-step estimator is usually preferred to the instrumental variables estimator and that it has properties which are very similar to those of the iterative Gauss-Newton estimator.

Efficient Estimation of the Geometric Distributed Lag Model
Some Monte Carlo Results on Small Sample Properties.

F.C. Palm
E. Vogelvang
D.A. Kodde*

October 1980.
Comments welcome

1. Introduction

In recent years, several asymptotically efficient two-step and iterative estimators for dynamic models with autocorrelated errors have been presented in the literature. Some results on the small sample properties of the two-step and iterative estimation procedures are also available. Among closely related Monte Carlo studies, we should like to mention the comparison of the finite sample properties of several estimators for the regression model with autoregressive errors by Rao and Griliches (1969) and for the Koyck (1954) distributed lag model by Morrison (1970) and Dhrymes (1971). Hatanaka (1974) presents an efficient two-step estimator for a single equation dynamic adjustment model with first order autoregressive errors and reports results of a simulation experiment. Hendry and Sbra (1977) investigate the small sample properties of instrumental variables estimators in a simultaneous equation framework with autoregressive errors. Harvey and McAvinchey (1979) compare the efficiency in small samples of various two-step and iterative estimation procedures for regression models with moving average errors.

In this paper, we report Monte Carlo results on instrumental variables, efficient two-step and iterative Gauss-Newton estimators of a Koyck distributed lag model with uncorrelated errors (model 1) and with first order autoregressive errors (model 2).

*Economische Faculteit, Vrije Universiteit, Amsterdam.
The authors wish to thank H.J. Blommestein for his help in programming the estimation methods. They are indebted to D.F. Hendry and A.C. Harvey for their helpful comments on an earlier version of this paper.

The distributed lag model with a Koyck scheme, perhaps the most widely used distributed lag model, is simple in the sense that it involves a small number of parameters. The parameter of the lag distribution can often be interpreted in terms of economic behavior such as adaptive expectation formation or partial adjustment. Still, the problems generally inherent in the estimation of distributed lag models are also present here, so that Koyck's model is a natural candidate for a simulation study.
In the last decade, dynamic specification analysis has received much attention in the econometric literature. As the different approaches to specification analysis require estimates of several alternative dynamic specifications, possibly arranged as a uniquely ordered sequence of restricted models, the demand for computionally convenient estimation methods with desirable small and large sample statistical properties has arisen. Usually one has to choose between consistent but inefficient or consistent and asymptotically efficient estimators, either iterative or not. The choice is usually based on criteria such as the computational costs involved, the small sample properties and the asymptotic efficiency.
In order to be able to offer some guidance for empirical work, we focus on the small sample properties of one estimator in each of the three classes of estimators, i.e. Liviatan's instrumental variables estimator, an efficient two-step and an iterative Gauss-Newton estimator. The latter is called a minimum chi-square estimator by Dhrymes (1971) [see also Dhrymes (1974)] , who shows that it becomes indistinguishable from the exact ML estimator in larger samples.
In section 2, we shortly present the models and the estimation procedures. A more detailed presentation of the estimation methods and their large sample properties can be found in e.g. Dhrymes, Klein and Steiglitz (1970), Harvey (1978) or in Palm (1978). In section 3, we describe the experiments. Section 4 contains the results of the simulations. They are summarized using response functions. Instead of generating a large number of runs for each experiment, we use the technique of control variates to increase the precision of the outcome of the simulations. In the last section, we draw some final conclusions.

2. The models and the estimation procedures

We analyze the geometric distributed lag model

$$y_t = \alpha_0 + \alpha_1 \sum_{i=0}^{\infty} \lambda^i x_{t-i} + u_t \quad , \quad t = 1, \dots T, \tag{2.1}$$

where $0 < \lambda < 1$ and x_t is independent of the error term $u_{t'}$, for all t and t', and T is the sample size.
We first consider the case where u_t is a white noise (model 1) with finite variance σ^2. Then we assume that u_t is generated by a first order autoregressive proces (model 2).
If the u_t's are independent and normally distributed, the likelihood function is

$$L(y, x, \alpha_0, \alpha_1, \lambda, \sigma^2) = (\sqrt{2\pi}\,\sigma)^{-T} \exp - \frac{1}{2\sigma^2} \sum_{t=1}^{T} (y_t - \alpha_0 - \alpha_1 x_t^*)^2 \, , \tag{2.2}$$

where the variable x_t^* is defined as

$$x_t^* = \sum_{i=0}^{\infty} \lambda^i x_{t-i} = \frac{1}{1-\lambda L} x_t \tag{2.3}$$

for a sequence of variables x_t with L being the lag-operator.

The first order conditions for a maximum of the log-likelihood function with respect to $\beta = (\alpha_0, \alpha_1, \lambda)'$ are given by

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma^2} X^{*\prime} u = 0 \quad , \tag{2.4}$$

with

$$X^{*\prime} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1^* & x_2^* & \dots & x_T^* \\ \alpha_1 x_0^{**} & \alpha_1 x_1^{**} & \dots & \alpha_1 x_{T-1}^{**} \end{bmatrix} , \tag{2.5}$$

$$x_t^{**} = \frac{1}{(1-\lambda L)^2} x_t \quad \text{and} \quad u = (u_1\, u_2 \dots u_T)' \, .$$

In the sequel we use the symbols "$\hat{}$" and "$\hat{\hat{}}$" to indicate that a variable is evaluated at the first and the second step parameter estimates respectively. The first order conditions (2.4) are nonlinear in the parameter vector β. We can solve them iteratively to obtain the maximum likelihood (ML) estimator. However, it is well-known (see e.g. Dhrymes & Taylor (1976)) that the following two-step estimator has the same asymptotic properties as the ML estimator of β

$$\hat{\hat{\beta}} = \hat{\beta} - \Gamma^{-1}(\hat{\beta}) \left. \frac{\partial \ln L}{\partial \beta} \right|_{\beta = \hat{\beta}} , \tag{2.6}$$

provided $\hat{\beta}$ is a consistent estimator of β such that $\sqrt{T}(\hat{\beta} - \beta_0)$, with β_0 being the true value of β, has some limiting distribution, and $\Gamma(\hat{\beta})$ is a non-singular matrix such that

$$\operatorname*{plim}_{T \to \infty} \frac{1}{T} \Gamma(\hat{\beta}) = \operatorname*{plim}_{T \to \infty} \frac{1}{T} \frac{\partial^2 \ln L(\beta_0)}{\partial \beta \; \partial \beta'} . \tag{2.7}$$

As the log-likelihood function is proportional to $u'u$, maximizing the likelihood function is equivalent to minimizing the sum of squares $u'u$. One way to implement (2.6), such that (2.7) is satisfied, is to compute one step of the Gauss-Newton algorithm starting with a consistent estimate of β_0, (see e.g. Palm (1978)). The formula for the Gauss-Newton algorithm is given by

$$\hat{\hat{\beta}} = \hat{\beta} - \left[\frac{\partial u}{\partial \beta} \; \frac{\partial u'}{\partial \beta} \right]^{-1} \left. \frac{\partial u}{\partial \beta} u \right|_{\beta = \hat{\beta}} \tag{2.8a}$$

$$= \hat{\beta} + (X^{*\prime} X^*)^{-1} X^{*\prime} u \Big|_{\beta = \hat{\beta}} \tag{2.8b}$$

as $\frac{\partial u}{\partial \beta} = - X^{*\prime}$ in (2.5). Iteration of (2.8) yields the nonlinear least squares estimator of β, which has the same asymptotic properties as the ML estimator. Whether the nonlinear least squares estimator is identical with a conditional or the exact ML estimator depends on the treatment of the initial values for the process x_t. Notice also that the difference between the two-step and the initial consistent estimator, $\hat{\hat{\beta}} - \hat{\beta}$, in (2.8) can be computed through an ordinary least squares regression of the residuals $\hat{u}$ on their partial derivatives with respect to β, both evaluated at $\hat{\beta}$. These derivatives can be computed analytically as in (2.5) or numerically (for numerically computed derivatives, see e.g. Harvey and McAvinchey (1979)). We use the analytical formula for the derivatives and compute the two-step estimator in (2.8) as follows:

1. Consistent parameter estimates are obtained by Liviatan's instrumental variables method applied to the transformed model

$$y_t = \alpha_0 (1 - \lambda) + \lambda y_{t-1} + \alpha_1 x_t + v_t , \quad t = 2, \ldots, T \tag{2.9}$$

with $v_t = u_t - \lambda u_{t-1}$, using x_{t-1} as an instrument for y_{t-1}. The restriction $0 < \lambda < 1$ is imposed on the estimate $\hat{\lambda}$. If $\hat{\lambda}$ lies outside the interval [.05 , .95] , it is fixed at the corresponding boundary value and the parameters α_0 and α_1 are estimated in a regression of $y_t - \hat{\lambda} y_{t-1}$ on x_t. The boundary values for $\hat{\lambda}$ were chosen after some experimentation with the model when $\lambda = .9$. For a boundary value very close to one and $\lambda = .9$, the iterative estimator of λ often has a cyclical behavior.

The variance of u_t is estimated by $\hat{\sigma}^2 = \frac{1}{T-4} \sum_{t=2}^{T} \hat{u}_t^2$, with $\hat{u}_2 = \hat{v}_2$ and $\hat{u}_t = \hat{\lambda} \hat{u}_{t-1} + \hat{v}_t$, $t = 3, \ldots, T$, where $\hat{v}_t$ is an instrumental variables residual.

2. In order to compute the two-step estimator in (2.8b) we rewrite the model (2.1) - after adding the same quantity to both sides of the equation - as

$$[y_t + \lambda \alpha_1 x^{**}_{t-1}] = \alpha_0 + \alpha_1 [x^*_t] + \lambda [\alpha_1 x^{**}_{t-1}] + u_t \tag{2.10}$$

It is straightforward to see that the two-step estimator of β in (2.8b) can be computed by ordinary least squares applied to the equation (2.10) after evaluation of the quantities between brackets at the consistent first step estimates.
Of course, there are many other ways to generate two-step estimators with the same asymptotic distribution as the ML estimator. Any matrix Γ satisfying the requirement (2.7) characterizes a two-step estimator, which is asymptotically equivalent to the ML estimator. For example the estimators proposed by Hannan (1965) and by Steiglitz and McBride (1965) have this property. The small sample properties of these estimators and of Liviatan's instrumental variables estimator for model 1 have been investigated by Morrison (1970).

We compute the two-step estimator of β in an OLS-regression of equation (2.10) for $t = 2, \ldots T$. The variables involved in the regressand and in the regressors of (2.10) are computed as

$$\hat{x}^*_t = x_t + \hat{\lambda} \hat{x}^*_{t-1} \qquad \text{and} \qquad \hat{x}^{**}_t = \hat{x}^*_t + \hat{\lambda} \hat{x}^{**}_{t-1}$$

with $\hat{x}_0^*$ and $\hat{x}_0^{**}$ being set equal to the sample mean of x_t and x_t^* respectively, divided by $1 - \hat{\lambda}$ (the process x_t is stationary). The estimate $\hat{\hat{\lambda}}$ has to lie inside the interval $[.05\ ,\ .95]$. Otherwise it is fixed at the corresponding boundary value and α_0 and α_1 are estimated in a regression of y_t on $\hat{x}_t^*$. Finally, the variance of u_t is estimated as in step 1 but using the residuals of step 2. When iterating the Gauss-Newton algorithm, we reestimate equation (2.10) by OLS after evaluation of the regressand and regressors between the brackets at the parameter estimates of the preceding step. The algorithm stops when convergence is achieved, i.e. the change in the estimates of α_1 and λ is smaller than .001, when the number of iterations is 100 or when the restriction on λ is violated for the second time.

In model 2, the disturbances u_t are generated by a first order autoregressive process

$$u_t = \rho\, u_{t-1} + \varepsilon_t \tag{2.11}$$

with $|\rho| < 1$, $\rho \neq \lambda$ and ε_t being a normally distributed white noise process with variance σ^2.

Equation (2.1) can be written as

$$y_t - \rho\, y_{t-1} = \alpha_0\,(1 - \rho) + \alpha_1\,(x_t^* - \rho\, x_{t-1}^*) + \varepsilon_t \tag{2.12}$$

and the two-step Gauss-Newton estimator for $\Theta = (\alpha_0\ \alpha_1\ \lambda\ \rho)'$ is given by

$$\hat{\hat{\Theta}} = \hat{\Theta} - \left[\frac{\partial\varepsilon}{\partial\Theta}\ \frac{\partial\varepsilon'}{\partial\Theta}\right]^{-1} \frac{\partial\varepsilon}{\partial\Theta}\,\varepsilon\,\Bigg|_{\Theta = \hat{\Theta}} \quad , \tag{2.13}$$

where $\hat{\Theta}$ is an initial consistent estimator of Θ, $\frac{\partial\varepsilon}{\partial\Theta}$ is the matrix of partial derivatives of the disturbance ε_t with respect to the elements in Θ

$$\frac{\partial\varepsilon}{\partial\Theta} = - \begin{bmatrix} 1-\rho & \dots & 1-\rho \\ x_1^* - \rho\, x_0^* & \dots & x_T^* - \rho\, x_{T-1}^* \\ \alpha_1(x_0^{**} - \rho\, x_{-1}^{**}) & \dots & \alpha_1(x_{T-1}^{**} - \rho\, x_{T-2}^{**}) \\ u_0 & \dots & u_{T-1} \end{bmatrix} \tag{2.14}$$

and $\varepsilon = (\varepsilon_1, \varepsilon_2\ \dots\ \varepsilon_T)'$ is the vector of disturbances.

The second right-hand-side term of (2.13) is evaluated at the consistent estimates $\hat{\Theta}$. The two-step estimator presented in (2.13) has the same asymptotic properties as the ML estimator, provided the requirements in (2.6) and (2.7) are satisfied. If we iterate the estimator (2.13) until convergence, we get the conditional ML estimator.

We compute the two-step estimator (2.13) as follows.

1. As for model 1, we estimate the parameters α_0, α_1 and λ consistently by instrumental variables applied to the transformed model (2.9) using x_{t-1} as an instrument for y_{t-1} and checking the restriction on λ. Then we compute $\hat{u}_t = \hat{v}_t + \hat{\lambda}\, \hat{u}_{t-1}$, $t = 3, \ldots T$, $\hat{u}_2 = \hat{v}_2$,

$$\hat{\rho} = \sum_{t=3}^{T} \hat{u}_t\, \hat{u}_{t-1} \Big/ \sum_{t=2}^{T} \hat{u}_t^2 \qquad \text{and}$$

$$\hat{\sigma}^2 = \frac{1}{T-4} \sum_{t=3}^{T} \hat{\varepsilon}_t^2 \ ,$$

where $$\hat{\varepsilon}_t = \hat{u}_t - \hat{\rho}\, \hat{u}_{t-1} \ .$$

2. Using expressions (2.12) and (2.14), it is straightforward to show that the two-step estimator in (2.13) can be computed by OLS applied to the following equation (which is obtained through adding the same terms to both sides of equation (2.12))

$$[y_t - \rho\, y_{t-1} + \lambda\, \alpha_1 (x^{**}_{t-1} - \rho\, x^{**}_{t-2}) + \rho\, u_{t-1}] =$$

$$= \alpha_0\, [1-\rho] + \alpha_1\, [x^*_t - \rho\, x^*_{t-1}] + \lambda[\alpha_1 (x^{**}_{t-1} - \rho\, x^{**}_{t-2})] +$$

$$+ \rho\, [u_{t-1}] + \varepsilon_t \quad , \quad t = 3, \ldots T \quad , \qquad (2.15)$$

after evaluation of the regressand and the regressors between brackets at consistent parameter estimator along the lines adopted for model 1. The restriction $.05 \leq \hat{\hat{\lambda}} \leq .95$ is also imposed in a similar way. The runs, for which the restriction $|\hat{\hat{\rho}}| \leq 1$ is not satisfied, are disregarded.

The latter restriction has been satisfied in most cases, although we do not use a block-diagonal matrix Γ in the two-step and iterative estimation procedure (for more details see e.g. Palm (1978)). When iterating the Gauss-Newton estimator for model 2, the algorithm stops if the change in the estimates of α_1 , λ and ρ is smaller than .001 or when the number of iterations is equal to 100. It also stops when the restriction on λ is violated for the second time.

Finally, notice that for both models we ignore the first observations. Whether this affects the conclusions about the finite sample properties, as has been found by Beach and MacKinnon (1978) for a linear regression model with autoregressive errors, has not been investigated.

3. The design of the experiments

The complete model used to generate the data is defined by the following

$$y_t = \alpha_0 + \alpha_1 \sum_{i=0}^{\infty} \lambda^i x_{t-i} + u_t , \quad 0 < \lambda < 1 \tag{3.1a}$$

$$u_t = \rho\, u_{t-1} + \varepsilon_t , \quad \rho \neq \lambda , \quad |\rho| < 1 \tag{3.1b}$$

$$\varepsilon_t \sim \mathrm{IN}\,(0, \sigma^2) \;\; \forall\, t , \tag{3.1c}$$

$$x_t = \gamma\, x_{t-1} + \eta_t , \quad 0 < \gamma < 1 , \tag{3.1d}$$

$$\eta_t \sim \mathrm{IN}\,(0, 10) \;\; \forall\, t , \tag{3.1e}$$

ε_t and $\eta_{t'}$ are independent for all t and t' .

The following parameter values are considered

$\alpha_0 = 50$, $\alpha_1 = .9$

$\lambda \in \{.3 , .6 , .9\}$

$\rho \in \{-.85 , -.5 , 0 , .5 , .85\}$

$\gamma \in \{0 , .7 , .95\}$

$\sigma \in \{5 , 10\}$.

These values cover the range of plausible values for the parameters and for the theoretical R^2 . The sample size T is equal to 30 and 60. The process for x_t is stationary and satisfies the Grenander conditions. For $\gamma = .95$, the spectrum for x_t approximately has the "typical shape of the spectrum of an economic variable". Using a trending x_t would imply a standardisation of the asymptotic distribution of the parameter estimate, which is different from $\sqrt{T}$.

Random samples of size $40 + T$ are generated from a uniform distribution. They are transformed into ε_t and η_t according to (3.1c) and (3.1e) using the probability integral theorem. The random variables u_t and x_t are generated according to (3.1b) and (3.1d) respectively, with $u_1 = \varepsilon_1 \sqrt{\frac{1}{1-\rho^2}}$ and $x_1 = \eta_1 \sqrt{\frac{1}{1-\gamma^2}}$.

Then, for a given set of parameter α_0 , α_1 and λ , sixty independent samples of size $40 + T$ for the variable y_t are generated using the model (3.1a), with $x_t = 0$ for $t \leq 0$. In order to guarantee the independence of y_t from the initial values of x_t , only the last T observations are used in the simulation study. As an alternative, we could have generated y_0 using its marginal density function implied by model (3.1) and the y_t's $t = 1, \ldots , T$ using equation (2.9).

4. The results of the simulations

For each of the sixty independent runs of an experiment, we estimate the parameters using Liviatan's instrumental variables (IV) method, the two-step (2S) and the iterative Gauss-Newton (IGN) estimation procedure as described in section 2. We compute and analyse the simulation mean and standard errors (SE) for these estimators. We do not investigate the existence of finite sample moments of the estimators. Rather we are interested in the relationships between simulation mean and SE's and the characteristics of the experiments. We model these relationship in response function equations and estimate them by OLS.
Furthermore, we focus our analysis on the appropriateness of large sample theory for finite sample situations. Possibly, the use of restricted estimators guarantees the existence of their finite sample moments.

In order to reduce the variance of the simulation results, we apply the technique of control variates (CV) to the outcome of the experiments (see e.g. Mikhail (1972, 1975)). For a more detailed description of this variance reduction technique, the reader is referred to e.g. Hendry and Srba (1977) and the references therein. In short, the basic idea can be presented as follows. Suppose that we want to simulate the finite sample mean (assumed to exist) of an estimator $\hat{\theta}$ of the parameter θ . We can compute the sample mean of the outcome $\hat{\theta}_j$ of m independent runs

$$\bar{\hat{\theta}} = \frac{1}{m} \sum_{j=1}^{m} \hat{\theta}_j \quad . \tag{4.1}$$

Consider now an alternative estimator $\bar{\theta}^o$ with known mean $E(\bar{\theta}^o)$. Then, the quantity $\tilde{\theta} = \bar{\hat{\theta}} - \bar{\theta}^o + E(\bar{\theta}^o)$ will have the same expectation as $\bar{\hat{\theta}}$. Its variance

$$\mathrm{var}(\tilde{\theta}) = \mathrm{var}(\bar{\hat{\theta}}) + \mathrm{var}(\bar{\theta}^o) - 2\,\mathrm{cov}(\bar{\hat{\theta}}, \bar{\theta}^o) \tag{4.2}$$

will be smaller than the variance of $\bar{\hat{\theta}}$, provided

$$2\,\mathrm{cov}(\bar{\hat{\theta}}, \bar{\theta}^o) > \mathrm{var}(\bar{\theta}^o) \quad . \tag{4.3}$$

The technique of CV's consists in choosing an estimator $\bar{\theta}^o$ (called CV) with known mean and satisfying (4.3) and to use $\tilde{\theta}$ instead of $\bar{\hat{\theta}}$ as an estimator of the unknown expectation of $\hat{\theta}$. In order to assure a high positive correlation between $\bar{\theta}^o$ and $\bar{\hat{\theta}}$, we derive the control variate $\bar{\theta}^o$ from the asymptotic distribution of $\hat{\theta}$. We choose $\bar{\theta}^o$ such that it has as finite sample distribution the large sample distribution of $\bar{\hat{\theta}}$.

For the IV estimator of $\Theta = (\beta' , \rho)'$ in model 2,

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'y \quad , \quad \hat{\rho}_{IV} = (\hat{u}'_{-1}\hat{u}_{-1})^{-1} \hat{u}'_{-1} \hat{u} \quad , \tag{4.4}$$

with $Z' = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_{T-1} \\ x_2 & x_3 & \dots & x_T \end{bmatrix}$, $X' = \begin{bmatrix} 1 & 1 & \dots & 1 \\ y_1 & y_2 & \dots & y_{T-1} \\ x_2 & x_3 & \dots & x_T \end{bmatrix}$

and $\hat{u}_{-1} = (\hat{u}_1 , \hat{u}_2 , \dots \hat{u}_{T-1})$ being the matrices of instruments and regressors and the vector of lagged residuals respectively, the CV's are given by

$$\beta^o_{IV} = E^{-1}(Z'X) Z'y \tag{4.5a}$$

and

$$\rho^o_{IV} = E^{-1}(u'_{-1} u_{-1}) u'_{-1} u = \frac{1 - \rho^2}{(T-1) \sigma^2} (u'_{-1} u) \quad . \tag{4.5b}$$

The control variate β^o_{IV} has as expectation β and as distribution the asymptotic distribution of $\hat{\beta}_{IV}$

$$\sqrt{T} (\hat{\beta}_{IV} - \beta) \overset{A}{\sim} N (0 , \Omega_{IV}) \quad , \tag{4.6}$$

with $\Omega_{IV} = T E^{-1}(Z'X) E(Z'VZ) E^{-1}(X'Z)$, where V is the covariance matrix of the vector $v = (v_2 , v_3 \dots v_T)'$.
The vector v is generated by an ARMA (1 , 1)-model $v_t = \frac{1 - \lambda L}{1 - \rho L} \varepsilon_t$, with autocovariances given by

$$E (v_t^2) = \frac{1 + \lambda^2 - 2 \rho \lambda}{1 - \rho^2} \sigma^2$$

$$E (v_t v_{t-1}) = \frac{(1 - \rho\lambda) (\rho - \lambda)}{1 - \rho^2} \sigma^2 \tag{4.7}$$

$$E (v_t v_{t-j}) = \rho E (v_t v_{t-j+1}) \quad , \quad j = 2, 3, \dots \; .$$

The control variate ρ^o_{IV} is centered at ρ and has as distribution the asymptotic distribution of $\hat{\rho}_{IV}$

$$\sqrt{T} (\hat{\rho}_{IV} - \rho) \overset{A}{\sim} N (0 , 1 - \rho^2) \quad . \tag{4.8}$$

Notice that $\hat{\rho}_{IV}$ and $\hat{\beta}_{IV}$ are independent in large samples.
The CV's given in (4.5) are expected to be almost perfectly correlated with the IV estimates in large samples. As the two-step and the iterative estimator have the same asymptotic distribution, we use the same CV's

$$\theta^o_{2S} = \theta^o_{IGN} = E^{-1}(P'P) P'y \quad , \tag{4.9}$$

where $P' = -\frac{\partial\varepsilon}{\partial\Theta}$ defined in (2.14) but for $t = 3, \ldots T$, so that it is of order $4 \times T-2$.
The mean of the control variates, $E(\Theta^o_{2S})$ is equal to the true parameter values. The finite sample distribution of Θ^o_{2S} is the same as the large sample distribution of the 2S-estimator

$$\sqrt{T}\ (\hat{\Theta}_{2S} - \Theta) \overset{A}{\sim} N\,(0\,, \sigma^2_\varepsilon\, T\, E^{-1}(P'P)) \quad . \tag{4.10}$$

The matrix $E(P'P)$ will be given in the appendix.
The CV's for model 1 are easily obtained from (4.5) and (4.9) by setting $\rho = 0$ and deleting the last column of P.

In the tables 1 - 3, we report the results of 12 experiments in detail. The values of the parameters and the sample size in these experiments are close to those often encountered in empirical econometric work.
In the columns 2, 7 and 13 of the tables 1 - 3, the simulation mean (M) for the IV, 2S and IGN estimators respectively of a parameter Θ_i is given

$$\bar{\hat{\Theta}}_i = \frac{1}{m} \sum_{j=1}^{m} \hat{\Theta}_{ij} \quad , \tag{4.11}$$

where $m = 60$ minus the number of times convergence is not achieved at step 100 or the restrictions on λ and/or ρ are not satisfied.
In columns 3, 8 and 14, the simulation standard errors (SSE) for the estimators are computed as

$$\left[\frac{1}{m-1} \sum_{j=1}^{m} (\hat{\Theta}_{ij} - \bar{\hat{\Theta}}_i)^2\right]^{\frac{1}{2}} . \tag{4.12}$$

In columns 4 and 9, the mean of the control variates for the IV and 2S estimator resp. (MCV) is given by

$$\bar{\Theta}^o_i = \frac{1}{m} \sum_{j=1}^{m} \Theta^o_{ij} \quad . \tag{4.13}$$

In columns 5 and 10, the standard deviation of the control variates (SDCV) are computed as

$$\left[\frac{1}{m-1} \sum_{j=1}^{m} (\Theta^o_{ij} - \bar{\Theta}^o_i)^2\right]^{\frac{1}{2}} \quad . \tag{4.14}$$

In columns 11 and 15, the square root of the mean of the variances of the estimators computed from the conventional formula for the estimated standard errors (ESE) is computed as

$$\left[\frac{1}{m-1} \sum_{j=1}^{m} DE_{ij}\right]^{\frac{1}{2}} \tag{4.15}$$

where DE_{ij} is the i-th diagonal element of $\hat{\sigma}_j^2 [\hat{P}_j' \hat{P}_j]^{-1}$ for run j, with $\hat{\sigma}_j^2$ and $\hat{P}_j$ being evaluated at the 2S and iterated estimates respectively. For the IV estimator, the appropriate formula Ω_{IV} for the estimated variance of $\hat{\beta}_{IV}$ is given in (4.6), with the moments replaced by their sample equivalents. As the formula is almost never used in empirical work, we have not computed ESE's for the IV estimator.

In columns 6 and 12 , the asymptotic standard errors (ASE) are equal to the square root of the i-th diagonal element of the covariance matrices in (4.6) and (4.10) divided by T. The reader can easily obtain a CV estimate of the finite sample bias of the IV estimator [2S , IGN] by substracting column 4 [9 , 9] from column 2 [7 , 13]. Similarly, a CV estimate of the variance of the IV estimator [2S , IGN] can be obtained by substracting the square of an element in column 5 [10 , 10] from that of the corresponding element in column 3 [8 , 14] and adding that of the asymptotic standard errors in column 6 [12 , 12]. Although a CV estimate of the variance is sometimes greater than the simulation variance, it is a more efficient estimate of the unknown variance. Notice also that for most of the experiments, the SSE's are closer to the ASE's than the ESE's. The variance of the estimates of α_0 is high and usually differs substantially from its asymptotic value. In those cases, the results for σ^2 are not very satisfactory either. Whether this is an indication of the non-existence of finite sample moments of the estimators or of possible multicollinearity has not been investigated. The bias of the 2S estimator of α_0 , for $\rho \neq 0$ and $T = 40$, is much greater than that of the IV or IGN estimator. Although we do not report additional results for the parameter α_0 , we should mention that they are not always satisfactory. In general, the results for the parameters α_1 , λ and ρ are satisfactory. The bias and the SE's of the 2S and IGN estimators for these parameters are very similar. The results in the tables do not indicate a dominance of IGN on the 2S estimator. For the 2S and IGN estimator in model 1, the SSE's are usually smaller than the ESE's. For model 2, both are fairly good - especially when $T = 60$ - , except for the parameter λ , for which the SSE is closer to the ASE than the ESE. The results in the tables 1 - 3 should give an overall picture of the finite sample properties of the three estimators considered. Still, they should not be carried over straightforwardly to other experiments.

Next, in order to give an impression of the gain in precision when using CV estimates for the mean of an estimator, we report in table 4 the ratio of the simulation variance over the CV variance for several selected experiments, i.e.

$$\mathrm{RVar} = \frac{\sum_{j=1}^{m} (\hat{\theta}_{ij} - \bar{\hat{\theta}}_{ij})^2}{\sum_{j=1}^{m} (\hat{\theta}_{ij} - \theta^o_{ij} - \bar{\hat{\theta}}_{ij} + \bar{\theta}^o_{ij})^2} \quad . \tag{4.16}$$

Except for high values of λ , there is usually a substantial reduction in the variance of the CV estimates, indicating that (4.3) is satisfied. When RVar = 2 , the gain in efficiency from the use of CV's is equal to that of doubling the number of runs. The response functions given in the tables 5 - 11 summarize the properties of the estimators for the experiments described in section 3. The tables 5 - 7 correspond to model 1. The response functions in tables 8 - 11 belong to model 2.

The response functions (RF) are estimated using 36 experiments for model 1 and 144 experiments for model 2. In each experiment the sixty independent samples for ε_t and η_t are reused, limiting thereby the computional costs at the price of some dependence. However under ergodicity, the results are not seriously affected. The functional form of response function is chosen after a detailed analysis of the plots of the outcome of the experiments as a function of the parameter values and the sample size T (see e.g. Figures 1 - 2). Thereby the results of the experiments were grouped according to the values of some parameters and the sample size.

We always impose the restriction on the RF specification that it should yield the asymptotic result for large values of T. As a dependent variable in the RF's for the bias, we use the standardized variable

$$B_i = \frac{\sqrt{m}\ (\bar{\hat{\theta}}_i - \theta_i)}{ASE_i} \tag{4.17}$$

for the simulation bias, and

$$BCV_i = \frac{\sqrt{m}\ (\bar{\hat{\theta}}_i - \bar{\theta}^o_i)}{ASE_i} \tag{4.18}$$

for the CV bias, where m is equal to the number of runs for which the restricted IGN estimator has converged.

Usually $m = 60$, but for values of λ and γ close to one, m migth be reduced to 40 . Notice that the RF's for the IV and 2S estimator are estimated from the results of the runs for which the IGN has converged.

The asymptotic distribution of the variable in (4.17) is $N(0, 1)$. A log-linear relationship between the SE's and the estimated residual variance and their asymptotic values (ASE and σ^2) is used. Additional terms depending on the remaining parameters and on T are needed in the specification in order to explain the variation of the SE's and the estimated residual variance over the experiments. Through the log-linear specification, we hope to achieve homoscedasticity (see e.g. Rao (1952)). For the 2S and the IGN estimator, the RF's of the SSE's and the ESE's are very similar. As the ESE's are more relevant to the empirical econometrician, we report RF's for them only. For the IV estimator, the RF's are estimated from the SSE-data. The CV estimates of the SE's are computed as

$$\text{SECV} = [\text{SSE}^2 - \text{SDCV}^2 + \text{ASE}^2]^{\frac{1}{2}} \tag{4.19a}$$

for the IV estimator, and

$$\text{SECV} = [\text{ESE}^2 - \text{SDCV}^2 + \text{ASE}^2]^{\frac{1}{2}} \tag{4.19b}$$

for the 2S and IGN estimator.

Usually the same specification for the RF's is retained whether direct simulation estimates or CV estimates are to be explained.
In the tables, the figures between brackets are standard errors. An explanatory variable written as $(x > c)$ takes the value 1 if x is larger than c and the value zero otherwise.

The RF's reported in the tables 5 - 11 have been used to predict the outcome of the independent experiments. In the tables 5 - 11, we give the value of

$$Q_i(l) = \frac{\sum_{j=1}^{l} (O_{ij} - P_{ij})^2}{S_i^2} , \tag{4.20}$$

where l is the number of independent experiments to be predicted, O_{ij} is the standardized outcome of experiment j for the parameter i, P_{ij} is the prediction from the response function and S_i^2 is the residual variance of the RF. Under the assumption that the RF is correctly specified and known, $Q_i(l)$ is approximately χ^2-distributed[1] with l degrees of freedom.
Alternatively, we also use the asymptotic $N(0, 1)$ distribution to predict the standardized outcome of an experiment. Under the assumption that the large sample distribution theory holds true for finite samples,

$$Q_{Ai}(l) = \sum_{j=1}^{l} O_{ij}^2 \tag{4.21}$$

is approximately χ^2-distributed with l degrees of freedom. Notice that the standardized CV estimates computed from (4.18) have a large sample variance, which is smaller than 1. Therefore the Q_{Ai} for the CV estimates should be rescaled in order to obtain a test-statistic which is approximately χ^2-distributed with l degrees of freedom.

---

1) This is not necessarily true for the predictions of the second order moments, as we use log-linear relationships.

The Q_i's and Q_{Ai}'s , for l equal to 4 and 8, are computed from the independent experiments given in the tables 1-3. As the outcome of the experiments for negative values of ρ exhibits great variability, we predict two additional independent experiments for $\rho = -.6$, $\gamma = .95$, $\lambda = .9$, $\sigma^2 = 10$ and $T = 40$ and 60 . The χ^2-values for these experiments are given in column 9 and 10 of the tables 8-11.[2)]

We shall now briefly draw some conclusions from the results in the tables 5-11. This should not dispense the reader from having a close look at the results themselves. Except for the standardized bias of the IV estimator, the form and the parameter values of the RF's for B_i and BCV_i are very similar. The residual standard deviation in the response functions for the bias decreases when the CV estimates are used. This does not happen for the RF's of the SE's. From the functional form of the response functions, it should be obvious that values of λ and γ close to the unit circle, of ρ close to -1 or a sample size T close to 30 heavily affect the finite sample properties of the three estimators considered in this paper. A similar conclusion has been drawn by Morrison (1970) for the small sample properties of Liviatan's IV estimator, a time domain version of Hannan's (1965) two-step estimator and of the iterative Steiglitz and McBride (1965) estimator in a geometric distributed lag model with uncorrelated errors.

The predictive power of the response functions is quite reasonable as is indicated by the values of the $Q_i(l)$'s . The RF for the bias of the IV estimator does not predict very well. The predictive performance of the large sample distribution theory in small sample situations is much less satisfactory. In comparison with estimated residual variance S_i^2 of the RF's , a large sample unit variance for the outcome of the experiments seems to be too small. This conclusion is not modified, if we predict the four experiments for $T = 60$ separately using the large sample $N(0,1)$ model. Notice also that the large sample theory implies testable restrictions for the response functions. For example, the coefficient of ln ASE should not be significantly different from one, while those of the remaining explanatory variables in the response functions for the SE's or for $\hat{\sigma}^2$ should not be significantly different from zero. This is not always confirmed by our analysis.
A major conclusion from the tables 5-11 is that the results for 2S and IGN are very similar, suggesting that for a sample of size $T \geq 30$, the applied econometrician can do without iterative estimation for the geometric distributed lag model.

2) The asterisk in the tables indicates that the χ^2-test is based on l-1 and l-2 predictions for model 1 and 2 respectively. For the excluded runs, the outcome for the CV estimate of the variance were negative. A negative R^2 as in table 8 can occur in models without constant term. In order to make the response functions compatible with the asymptotic theory, we do not include a constant

5. Some final conclusions

In this paper we have investigated the finite sample behavior of three estimators for the geometric distributed lag model using Monte Carlo experiments. We have tried to increase the precision of the outcome of the experiments through the use of control variates derived from the asymptotic distribution of the estimators. While the CV's yield a reduction of the variance of the results, the form and the point estimates of the RF's for the CV estimates of the bias and SE's are quite similar to those for the direct simulation results. Certainly, the gain in precision is lower than the increase in precision obtained by e.g. Hendry and Srba (1977). However, a major difference between their models and ours is the nonlinearity in the parameter λ of our model.
An important conclusion from our study is that the small sample ($T \geq 30$) properties of the two-step and of the iterative Gauss-Newton estimator are very similar, suggesting that it will in general be sufficient to compute an efficient two-step estimator.

Our results do not give much evidence about the possible non-existence of finite sample moments of the three estimators that we have considered. Perhaps the restrictions imposed on λ and ρ assure the existence of moments in finite samples. Possibly, we obtained good estimates of the Nagar approximations to the moments (see Sargan (1978)). Finally, as the response functions presented in this paper yield the asymptotic result for large T, they enable us to answer questions such as "What is a large sample?", "How large is large?". That the answer to these questions depends on the true parameter values (or what one might think as being the true parameter values) should be obvious.

Appendix

We shall give the elements of the matrix $E(P'P) = A$ as functions of the parameters of the model (3.1). Summation goes from $t = 3$ to T.
Denoting the i,j th element of the symmetric matrix A by a_{ij}, we have

$$a_{11} = (T-2)(1-\rho)^2$$

$$a_{12} = E[\Sigma (x_t^* - \rho x_{t-1}^*)(1-\rho)] = 0$$

$$a_{13} = 0$$

$$a_{14} = 0$$

$$a_{22} = E[\Sigma (x_t^* - \rho x_{t-1}^*)^2] = (T-2)[(1+\rho^2) E(x_t^{*2}) - 2\rho E(x_t^* x_{t-1}^*)]$$

$$a_{23} = E[\Sigma \alpha_1 (x_t^* - \rho x_{t-1}^*)(x_{t-1}^{**} - \rho x_{t-2}^{**})]$$

$$= (T-2)\alpha_1 [(1+\rho^2) E(x_t^* x_{t-1}^{**}) - \rho E(x_t^* x_t^{**}) - \rho E(x_t^* x_{t-2}^{**})]$$

$$a_{24} = E[\Sigma (x_t^* - \rho x_{t-1}^*) u_{t-1}] = 0$$

$$a_{33} = E[\alpha_1^2 \Sigma (x_{t-1}^{**} - \rho x_{t-2}^{**})^2]$$

$$= (T-2)\alpha_1^2 [(1+\rho^2) E(x_t^{**2}) - 2\rho E(x_t^{**} x_{t-1}^{**})]$$

$$a_{34} = E[\alpha_1 \Sigma (x_{t-1}^{**} - \rho x_{t-2}^{**}) u_{t-1}]$$

$$= 0$$

$$a_{44} = E[\Sigma u_{t-1}^2] = \frac{(T-2)\sigma^2}{1-\rho^2} .$$

Next we must express the second order moment of x_t^* and x_t^{**} as functions of the parameters λ, γ and σ_η^2. Notice that x_t^* and x_t^{**} are generated by a second and third order autoregressive process respectively with mean zero

$$x_t^* = \frac{1}{(1-\lambda L)(1-\gamma L)} \eta_t , \quad x_t^{**} = \frac{1}{(1-\lambda L)^2(1-\gamma L)} \eta_t .$$

The variance of the AR(2) process x_t is given by

$$E(x_t^{*2}) = \frac{\sigma_\eta^2 (1+\gamma\lambda)}{1 + (\gamma\lambda)^2 - \gamma\lambda - \gamma^2 - \lambda^2 + \gamma^3\lambda + \gamma\lambda^3 - (\gamma\lambda)^3} .$$

The first order autocovariance is

$$E(x_t^* x_{t-1}^*) = \frac{\sigma_\eta^2 (\gamma+\lambda)}{1 + (\gamma\lambda)^2 - \gamma\lambda - \gamma^2 - \lambda^2 + \gamma^3\lambda + \gamma\lambda^3 - (\gamma\lambda)^3} .$$

The variance of x_t^{**} is

$$E\,(x_t^{**2}) = \frac{\sigma_\eta^2}{1 - \psi_1\,\rho_1 - \psi_2\,\rho_2 - \psi_3\,\rho_3}$$

with

$$\psi_1 = \gamma + 2\lambda$$

$$\psi_2 = -\,(\lambda^2 + 2\gamma\lambda)$$

$$\psi_3 = \lambda^2\gamma$$

$$\rho_1 = \frac{\psi_1}{1-\psi_2} + \frac{\psi_3}{1-\psi_2}\left[\frac{\psi_1^2 + \psi_1\,\psi_3 + \psi_2 - \psi_2^2}{1 - \psi_2 - \psi_3(\psi_1 + \psi_3)}\right]$$

$$\rho_2 = \frac{\psi_1^2 + \psi_1\,\psi_3 + \psi_2 - \psi_2^2}{1 - \psi_2 - \psi_3(\psi_1 + \psi_3)}$$

$$\rho_3 = \psi_1\,\rho_2 + \psi_2\,\rho_1 + \psi_3 \quad .$$

The first order autocovariance of x_t^{**} is

$$E\,(x_t^{**}\,x_{t-1}^{**}) = \rho_1\,E\,(x_t^{**2}) \quad .$$

The cross-covariances are

$$E\,(x_t^{*}\,x_t^{**}) = \frac{B_1}{1-\lambda^2} + \frac{B_2}{1-\gamma\lambda}$$

where

$$B_1 = \frac{\sigma_\eta^2\,\lambda\,(1-\gamma^2)}{(\lambda-\gamma)[1 + (\gamma\lambda)^2 - \gamma\lambda - \gamma^2 - \lambda^2 + \gamma^3\lambda + \gamma\lambda^3 - (\gamma\lambda)^3]}$$

$$B_2 = \frac{-\,B_1\gamma\,(1-\gamma^2)}{\lambda\,(1-\gamma^2)}$$

$$E\,(x_t^{*}\,x_{t-1}^{**}) = \frac{B_1\,\lambda}{1-\lambda^2} + \frac{B_2\,\gamma}{1-\gamma\lambda} \quad ,$$

$$E\,(x_t^{*}\,x_{t-2}^{**}) = \frac{B_1\,\lambda^2}{1-\lambda^2} + \frac{B_2\,\gamma^2}{1-\lambda\gamma} \quad .$$

Finally notice that the matrix $E\,(Z'X)$ for the control variates of the IV estimator is obtained in a similar way.

References

Beach, Ch.M., and J.G. MacKinnon (1978): "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", Econometrica, 46, 51-58.

Dhrymes, P.J. (1971): Distributed Lags: Problems of Estimation and Formulation, San Francisco, Holden-Day Inc.

Dhrymes, P.J. (1974): "A note on an Efficient Two-Step Estimator", Journal of Econometrics, 2, 301-304.

Dhrymes, P.J., Klein, L.R., and K. Steiglitz (1970): "Estimation of Distributed Lag Models", International Economic Review, 11, 235-250.

Dhrymes, P.J., and J.B. Taylor (1976): "On an Efficient Two-Step Estimator for Dynamic Simultaneous Equations Models with Autoregressive Errors", International Economic Review, 17, 362-376.

Hannan, E.J. (1965): "The Estimation of Relationships Involving Distributed Lags", Econometrica, 33, 206-224.

Harvey, A.C. (1978): "Two-Step and Full Maximum Likelihood Estimators in Distributed Lag and Dynamic Adjustment Models", mimeographed, U. of British Columbia, Vancouver.

Harvey, A.C., and I.D. McAvinchey (1979): "On the Relative Efficiency of Various Estimators of Regression Models with Moving Average Disturbances", mimeographed.

Hatanaka, M. (1974): "An Efficient Two-Step Estimator for the Dynamic Adjustment Model with Autoregressive Errors", Journal of Econometrics, 2, 199-220.

Hendry, D.F. (1977): "Monte Carlo Methods for Investigating Finite Sample Distributions", L.S.E., mimeographed.

Hendry, D.F., and F. Srba (1977): "The Properties of Autoregressive Instrumental Variables Estimators in Dynamic Systems", Econometrica, 45, 969-990.

Koyck, L.M. (1954): Distributed Lags and Investment Analysis, Amsterdam, North-Holland Publishing Company.

Mikhail, W.M. (1972): "Simulating the Small Sample Properties of Econometric Estimators", Journal of the American Statistical Association, 67, 620-624.

Mikhail, W.M. (1975): "A Comparative Monte Carlo Study of the Properties of Econometric Estimators", Journal of the American Statistical Association, 70, 94-104.

Morrison, Jr., J.L. (1970): "Small Sample Properties of Selected Distributed Lag Estimators", International Economic Review, 11, 13-23.

Palm, F.C. (1978): "On efficient Two-Step Estimation of a Simple Distributed Lag Model", Vrije Universiteit, Amsterdam, Economische Faculteit, Diskussienota 2.

Rao, C.R. (1952): Advanced Statistical Methods in Biometric Research. New York, John Wiley and Sons.

Rao, P., and Z. Griliches (1969): "Small-Sample Properties of Several Two-Stage Regression Methods in the Context of Autocorrelated Errors", Journal of the American Statistical Association, 64, 253-272.

Sargan, J.D. (1978): "On Monte Carlo Estimator of Moments which are Infinite, mimeographed L.S.E.

Steiglitz, K., and L.B. McBride (1965): "A Technique for Identification of Linear Systems", IEEE Transactions on Automatic Control, AC-10, 461-464.

Table 1. Simulation Results for Some Selected Experiments with Model 1

($\alpha_0 = 50$, $\alpha_1 = .9$, $\gamma = .85$, $\sigma^2 = 10$) .

	IV					2S						IGN		
$\lambda=.7$, T=40	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	50.005	2.301	49.193	4.165	3.675	49.936	1.667	49.988	.466	1.178	.506	49.758	1.330	1.030
α_1	.8677	.1749	.8644	.1002	.2366	.9125	.1127	.8954	.0642	.0818	.0638	.9591	.0975	.0846
λ	.71539	.10049	.71616	.08378	.10736	.70710	.03887	.70284	.02107	.03847	.02320	.69180	.03171	.04210
σ^2	15.093	3.407				12.499	3.532					12.224	3.423	
$\lambda=.9$, T=40	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	49.501	9.428	50.250	2.272	4.756	47.851	18.410	50.046	.536	5.390	.506	46.630	11.760	4.046
α_1	.9341	.1473	.9239	.1760	.2898	1.0038	.2499	.9004	.0150	.0776	.0231	1.0461	.2531	.0794
λ	.88990	.09145	.89498	.04145	.07569	.89544	.03488	.90031	.00228	.01866	.00308	.89112	.03213	.02040
σ^2	23.663	14.210				42.205	37.704					38.452	35.239	
$\lambda=.7$, T=60	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	50.791	5.698	[illegible]	3.785	4.776	48.463	12.156	49.974	.401	3.836	.412	50.067	.767	.658
α_1	.9476	.1712	.9459	.1693	.1940	.9225	.0839	.9017	.0569	.0696	.0519	.9379	.0738	.0647
λ	.67728	.09156	.67865	.07529	.08729	.69806	.03444	.69940	.02132	.02889	.01886	.69224	.02811	.02738
σ^2	12.535	3.493				11.871	2.706					11.694	2.739	
$\lambda=.9$, T=60	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	48.695	8.924	50.541	2.630	3.675	50.094	15.054	49.968	.411	2.652	.412	49.892	10.343	2.488
α_1	.9625	.1587	.9124	.1978	.2350	.9571	.2140	.9020	.0150	.0707	.0188	1.0065	.2023	.0580
λ	.87526	.06865	.87706	.05089	.06193	.90785	.02548	.89970	.00195	.01588	.00250	.89748	.01852	.01097
σ^2	38.678	29.175				49.140	43.768					45.533	42.699	

Table 2. Simulation Results for Some Selected Experiments with Model 2

($\alpha_0 = 50$, $\alpha_1 = .9$, $\gamma = .85$, $\sigma^2 = 10$, $\rho = .4$)

γ=.7, T=40, ρ=.8	IV M	IV SSE	IV MCV	IV SDCV	IV ASE	2S M	2S SSE	2S MCV	2S SDCV	2S ESE	2S ASE	IGN M	IGN SSE	IGN ESE
α_0	50.876	11.337	[illegible]	3.319	4.321	46.367	22.327	51.193	62.127	15.445	2.565	50.345	5.761	3.508
α_1	.8957	.1375	.8957	.1212	.1649	.9266	.1456	.9924	.1051	.1445	.1389	.9208	.1492	.1447
λ	.70761	[illegible]	.70733	.06326	.07495	.69019	.08171	.70762	.05459	.07092	.06039	.69828	.08049	.07407
ρ	.66911	[illegible]	[illegible]	[illegible]	.09467	.64577	.14474	.74238	.07238	.13269	.09733	.64013	.15377	.13409
σ^2	9.913	2.339				10.111	2.349					10.244	2.466	

λ=.9, T=40, ρ=.8	IV M	IV SSE	IV MCV	IV SDCV	IV ASE	2S M	2S SSE	2S MCV	2S SDCV	2S ESE	2S ASE	IGN M	IGN SSE	IGN ESE
α_0	48.846	13.379	49.040	1.781	2.428	59.551	28.179	57.492	64.131	18.306	2.565	51.654	7.675	7.415
α_1	.9006	.1112	.8901	.1182	.1646	.9429	.1192	.9109	.0701	.1126	.0850	.9685	.1547	.1055
λ	.90205	.03130	.90772	.03525	.04297	.88346	.03460	.89943	.01304	.03507	.01251	.88649	.03796	.03270
ρ	.75950	.12804	.80156	.08662	.09487	.66367	.15296	.75831	.09891	.12902	.09733	.62370	.23959	.19342
σ^2	10.278	2.526				10.302	2.714					12.339	9.040	

λ=.7, T=60, ρ=.8	IV M	IV SSE	IV MCV	IV SDCV	IV ASE	2S M	2S SSE	2S MCV	2S SDCV	2S ESE	2S ASE	IGN M	IGN SSE	IGN ESE
α_0	50.090	2.534	50.216	3.144	3.523	49.972	2.658	49.468	53.718	2.091	2.076	49.965	2.979	2.145
α_1	.9082	.1642	.9024	.1550	.1238	.8966	.1456	.8880	.1265	.1128	.1124	.8936	.1419	.1115
λ	.70087	.08517	.69700	.06436	.06621	.70321	.08221	.70351	.05253	.05156	.04888	.70504	.08634	.05295
ρ	.70770	.09420	[illegible]	.08569	.07746	.70246	.10691	.75565	.05999	.09779	.07878	.70107	.10934	.09797
σ^2	9.534	1.595				9.661	1.574					9.665	1.555	

λ=.9, T=60, ρ=.8	IV M	IV SSE	IV MCV	IV SDCV	IV ASE	2S M	2S SSE	2S MCV	2S SDCV	2S ESE	2S ASE	IGN M	IGN SSE	IGN ESE
α_0	[illegible]	[illegible]	[illegible]	[illegible]	1.952	52.907	10.989	49.300	54.649	8.231	2.076	51.507	5.784	7.174
α_1	[illegible]	.1171	[illegible]	.1216	.1339	.9404	.1282	.9061	.0720	.0910	.0698	.9507	.1201	.0838
λ	.89425	.03773	[illegible]	.03009	.03494	.89786	.02417	.89979	.00948	.02372	.01013	.89889	.02073	.02177
ρ	[illegible]	[illegible]	[illegible]	.08705	.07746	.69699	.15675	.75140	.07681	.10618	.07878	.65389	.20697	.14982
σ^2	10.215	[illegible]				10.394	2.641					12.434	5.835	

As the ASE's for ρ for the IV and the 2S-estimators are divided by the number of observations used in the estimation, T-1 and T-2 respectively, the ASE's for the 2S-estimator are greater than those for the IV estimator.

Table 3. Simulation Results for Some Selected Experiments with Model 2

$(\alpha_0 = 50 \ , \ \alpha_1 = .9 \ , \ \gamma = .85 \ , \ \sigma^2 = 10 \ , \ \rho = .8)$

$\lambda=.7, T=40$	IV					2S						IGN		
$\rho = .4$	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	50.009	1.941	49.497	3.493	4.257	49.688	2.100	50.063	5.026	1.820	.855	49.760	1.326	1.441
α_1	.3749	.1524	.9736	.1586	.1773	.9288	.1009	.8900	.0968	.1165	.0964	.9310	.0953	.1128
λ	.70909	.08603	.70969	.07038	.07978	.68998	.04345	.70377	.03195	.05500	.03566	.69624	.04102	.05442
ρ	.39459	.17761	.37479	.14673	.14491	.30222	.17945	.31556	.14331	.16414	.14868	.28477	.17918	.16424
σ^2	10.235	4.747				10.227	2.563					10.227	2.553	

$\lambda=.9, T=40$	IV					2S						IGN		
$\rho = .4$	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	48.755	3.555	50.093	1.915	3.172	51.975	16.514	50.698	4.557	8.234	.955	49.762	8.457	4.134
α_1	.9236	.2109	.9121	.1510	.2076	.9730	.1495	.8998	.0257	.0948	.0380	.9950	.1694	.0805
λ	.89272	.04195	.89818	.03789	.05425	.89375	.03333	.90058	.00368	.02696	.00510	.89290	.02463	.02100
ρ	.53732	.16842	.37954	.14569	.14491	.33946	.22740	.33580	.17956	.18071	.14868	.26117	.28615	.21287
σ^2	11.479	2.351				13.122	6.774					16.869	13.600	

$\lambda=.7, T=60$	IV					2S						IGN		
$\rho = .4$	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	50.640	4.651	50.733	3.281	3.460	45.122	37.826	49.700	4.021	10.877	.692	50.053	.860	.952
α_1	.9309	.1575	.9339	.1476	.1441	.9285	.1124	.9074	.0681	.0949	.0780	.9260	.1018	.0878
λ	.68304	.07749	.68417	.06504	.06486	.68677	.04741	.69718	.03358	.04168	.02886	.69342	.04259	.03841
ρ	.38905	.12211	.38143	.11496	.11832	.32857	.13250	.33536	.11114	.12760	.12035	.31667	.13155	.12800
σ^2	10.542	2.199				10.435	1.912					10.406	1.952	

$\lambda=.9, T=60$	IV					2S						IGN		
$\rho = .4$	M	SSE	MCV	SDCV	ASE	M	SSE	MCV	SDCV	ESE	ASE	M	SSE	ESE
α_0	49.865	7.059	50.135	1.719	2.575	52.183	13.864	50.102	4.132	7.220	.692	51.176	5.755	2.633
α_1	.9345	.1481	.9137	.1357	.1480	.9554	.1216	.9027	.0233	.0868	.0308	.9828	.1284	.0604
λ	.89314	.04417	.89737	.03471	.04411	.89228	.03250	.89945	.00339	.02176	.00413	.89626	.01579	.01465
ρ	.58553	.20444	.35321	.12091	.11832	.40398	.19768	.30886	.15912	.14004	.12035	.25708	.26675	.22258
σ^2	12.356	3.862				13.272	5.776					18.755	13.751	

Table 4. Efficiency Gains for the Bias through the Use of Control Variates, defined as the Ratio of Variances in (4.16).

(γ = .95 , σ^2 = 10)

ρ	λ	T	θ_i	IV	2S	IGN
0	.3	30	α_1	5.12	4.48	4.19
		30	λ	5.60	2.89	2.99
		60	α_1	8.99	5.48	5.85
		60	λ	9.16	4.52	4.96
	.9	30	α_1	3.92	1.01	1.01
		30	λ	2.08	1.00	1.00
		60	α_1	3.97	1.00	1.00
		60	λ	3.54	.99	.99
.5	.3	30	α_1	3.13	3.78	3.71
		30	λ	2.05	1.47	1.49
		30	ρ	1.56	2.71	2.64
		60	α_1	6.30	6.87	6.55
		60	λ	6.50	6.05	5.39
		60	ρ	2.81	4.01	2.75
	.9	30	α_1	3.79	1.01	1.02
		30	λ	2.13	1.00	1.02
		30	ρ	1.06	1.15	.76
		60	α_1	4.12	1.02	1.01
		60	λ	2.87	.96	.99
		60	ρ	.89	1.22	.98
.85	.3	30	α_1	1.45	3.24	3.30
		30	λ	1.01	1.41	1.73
		30	ρ	.87	1.38	1.38
		60	α_1	2.52	4.65	4.71
		60	λ	3.45	2.75	2.86
		60	ρ	1.78	2.82	2.82
	.9	30	α_1	2.91	1.10	1.11
		30	λ	2.09	.94	.93
		30	ρ	.64	1.03	1.02
		60	α_1	3.95	1.15	1.07
		60	λ	2.77	1.04	1.02
		60	ρ	.66	1.47	1.06

Table 5. Response Functions for α_1 in Model 1

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(4)$	$Q_{Ai}(4)$
IV	B	$\frac{1}{\sqrt{T}}[-\underset{(3.03)}{11.10} + \underset{(3.09)}{10.25}\,\lambda + \underset{(3.30)}{1.37}\,\sigma^2]$	.500	1.81	.72	15.81	9.30
2S	B	$\frac{1}{\sqrt{T}}[-\underset{(13.46)}{4.90} + \underset{(.23)}{3.34}\,\frac{\lambda\gamma}{(1-\lambda)(1-\gamma)}]$	.860	1.54	11.439	.465	1,562.36
IGN	B	$\frac{1}{\sqrt{T}}[\underset{(18.96)}{2.86} + \underset{(.325)}{4.08}\,\frac{\lambda\gamma}{(1-\lambda)(1-\gamma)}]$	.820	1.98	16.115	1.51	3,947.5
IV	BCV	$\frac{1}{\sqrt{T}}[\underset{(1.46)}{10.94} - \underset{(1.49)}{1.65}\,\lambda - \underset{(.15)}{.75}\,\sigma^2]$	.460	1.15	1.15	7.03	.18
2S	BCV	$\frac{1}{\sqrt{T}}[-\underset{(13.18)}{6.51} + \underset{(.22)}{3.33}\,\frac{\lambda\gamma}{(1-\lambda)(1-\gamma)}]$	.864	1.58	11.21	12.42	1,516.48
IGN	BCV	$\frac{1}{\sqrt{T}}[\underset{(18.66)}{1.24} + \underset{(.32)}{4.07}\,\frac{\lambda\gamma}{(1-\lambda)(1-\gamma)}]$	.824	2.02	15.86	15.70	3,872.03
IV	ln SSE	$\underset{(.018)}{.93}\,\ln ASE + \underset{(2.99)}{29.82}\,\frac{1}{T} - \underset{(.23)}{2.75}\,\frac{\sigma^2}{T} - \underset{(2.19)}{20.08}\,\frac{\lambda}{T}$	.883	1.76	.082	30.85	.73
2S	ln SE	$\underset{(.019)}{.886}\,\ln ASE - \underset{(.125)}{.217}\,(\frac{1-\lambda}{T-29}) + \underset{(6.32)}{69.06}\,\frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.636	2.49	.207	2.25	3.38
IGN	ln SE	$\underset{(.215)}{.890}\,\ln ASE - \underset{(.14)}{.211}\,(\frac{1-\lambda}{T-29}) + \underset{(7.3)}{67.65}\,\frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.505	2.21	.239	.693	2.93
IV	ln SSE CV	$\underset{(.014)}{.98}\,\ln ASE + \underset{(1.72)}{7.30}\,\frac{1}{T} - \underset{(1.94)}{9.44}\,\frac{\lambda}{T}$	.970	1.57	.073	442.71*	1.03*
2S	ln SE CV	$\underset{(.020)}{.89}\,\ln ASE - \underset{(.14)}{.20}\,(\frac{1-\lambda}{T-29}) + \underset{(6.94)}{69.90}\,\frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.622	2.13	.23	63.82	3.32
IGN	ln SE CV	$\underset{(.022)}{.89}\,\ln ASE - \underset{(.15)}{.20}\,(\frac{1-\lambda}{T-29}) + \underset{(7.67)}{68.90}\,\frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.518	2.01	.25	46.83	2.86

Table 6. Response Functions for λ in Model 1

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(4)$	$Q_{Ai}(4)$
IV	B	$\frac{1}{\sqrt{T}}\left[\underset{(3.74)}{29.93} - \underset{(.20)}{2.48} \left(\frac{1}{1-\lambda}\right) - \underset{(.44)}{2.51}\ \sigma^2 \right]$	.760	2.00	1.05	9.15	15.33
2S	B	$\frac{1}{\sqrt{T}}\left[\underset{(14.47)}{18.70} - \underset{(.248)}{2.46} \frac{\lambda\gamma}{(1-\lambda)(1-\gamma)} \right]$	.747	1.55	12.296	9.82	686.41
IGN	B	$\frac{1}{\sqrt{T}}\left[\underset{(8.87)}{14.17} - \underset{(.152)}{3.18} \frac{\lambda\gamma}{(1-\lambda)(1-\gamma)} \right]$	.928	1.62	7.537	2.5	492.51
IV	BCV	$\frac{1}{\sqrt{T}}\left[\underset{(1.89)}{5.60} - \underset{(.15)}{1.60} \left(\frac{1}{1-\lambda}\right) - \underset{(.22)}{.18}\ \sigma^2 \right]$	.790	2.19	.53	54.54	2.37
2S	BCV	$\frac{1}{\sqrt{T}}\left[\underset{(14.53)}{18.41} - \underset{(.25)}{2.43} \frac{\lambda\gamma}{(1-\lambda)(1-\gamma)} \right]$	.741	1.59	12.35	4.51	743.91
IGN	BCV	$\frac{1}{\sqrt{T}}\left[\underset{(8.75)}{13.89} - \underset{(.15)}{3.15} \frac{\lambda\gamma}{(1-\lambda)(1-\gamma)} \right]$	.929	1.64	7.43	8.92	513.12
IV	ln SE	$\underset{(.038)}{.86} \ln ASE + \underset{(7.28)}{13.14} \left(\frac{1}{T}\right) - \underset{(6.15)}{10.04} \left(\frac{\lambda}{T}\right) - \underset{(.63)}{2.27} \frac{\sigma^2}{T}$	.710	1.75	.24	1.69	.15
2S	ln SE	$-\underset{(.232)}{.413} \left(\frac{1-\lambda}{T-29}\right) + \underset{(.026)}{.799} \ln ASE + \underset{(12.84)}{69.73} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.721	1.88	.401	2.09	7.11
IGN	ln SE	$-\underset{(.144)}{.211} \left(\frac{1-\lambda}{T-29}\right) + \underset{(.022)}{.890} \ln ASE + \underset{(7.30)}{67.64} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.505	2.21	.311	8.01	6.23
IV	ln SSECV	$\underset{(.027)}{.91} \ln ASE - \underset{(3.53)}{5.67} \left(\frac{1}{T}\right) + \underset{(4.58)}{3.99} \left(\frac{\lambda}{T}\right)$	.90	1.67	.18	12.73*	.13*
2S	ln SE CV	$-\underset{(1.23)}{.45} \left(\frac{1-\lambda}{T-29}\right) + \underset{(.027)}{.80} \ln ASE + \underset{(13.12)}{70.08} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.710	1.82	.41	42.30	7.08
IGN	ln SE CV	$-\underset{(.19)}{.32} \left(\frac{1-\lambda}{T-29}\right) + \underset{(.022)}{.83} \ln ASE + \underset{(10.54)}{81.67} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.817	2.22	.33	57.48	6.23

Table 7. Response Functions for σ^2 in Model 1

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(4)$	$Q_{Ai}(4)$
IV	$\ln \hat{\sigma}^2$	$\underset{(.074)}{1.15} \ln \sigma^2 + \underset{(10.85)}{24.98} (\frac{\lambda}{T}) - \underset{(7.83)}{17.37} \frac{1}{T} + \underset{(12.13)}{44.7} \frac{\lambda\gamma^4}{T}$	.704	2.19	.314	4.16	2.79
2S	$\ln \hat{\sigma}^2$	$\underset{(.105)}{1.19} \ln \sigma^2 - \underset{(8.07)}{10.74} \frac{1}{T} + \underset{(11.46)}{102.64} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.726	2.33	.448	.864	4.69
IGN	$\ln \hat{\sigma}^2$	$\underset{(.105)}{1.18} \ln \sigma^2 - \underset{(8.06)}{10.77} \frac{1}{T} + \underset{(11.45)}{98.26} \frac{(\lambda > .7)(\lambda\gamma)^2}{T}$	.711	2.31	.448	.841	4.18

Table 8. Response Functions for α_1 in Model 2

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(8)$	$Q_{Ai}(8)$	$Q_i(2)$	$Q_{Ai}(2)$
IV	B	$\frac{1}{\sqrt{T}}[.44\ \sigma^2 - 2.36\ \lambda + .13\ \rho + 9.68\ \lambda^2\gamma]$ (.11) (1.67) (.58) (1.89)	.27	1.39	.76	16.41	8.05	.77	2.05
2S	B	$\frac{1}{\sqrt{T}}[-46.2 + 95.37\ \lambda^2\rho + 198.33\ \lambda^3(1-\rho)(1+\gamma^2)]$ (11.25) (30.37) (16.63)	.573	1.40	15.20	.88	350.04	37.81	15,392
IGN	B	$\frac{1}{\sqrt{T}}[-55.75 + 119.24\ \lambda^2\rho + 268.82\ \lambda^3(1-\rho)(1+\gamma^2)]$ (15.20) (41.04) (22.49)	.576	1.09	20.54	1.00	672.38	7.97	17,863
IV	BCV	$\frac{1}{\sqrt{T}}[.08\ \sigma^2 + 1.37\ \lambda + 2.84\ \rho + 3.44\ \lambda^2\gamma]$ (.06) (.94) (.32) (1.06)	.32	1.19	.43	15.40	1.67	.69	0.25
2S	BCV	$\frac{1}{\sqrt{T}}[-43.29 + 117.69\ \lambda^2\gamma + 132.67\ \lambda^3(1-\rho)(1+\gamma^2)]$ (10.91) (38.11) (15.34)	.564	1.40	15.31	.45	327.01	45.14	15,559
IGN	BCV	$\frac{1}{\sqrt{T}}[-53.73 + 161.31\ \lambda^2\gamma + 183.84\ \lambda^3\ (1-\rho)(1+\gamma^2)]$ (14.64) (51.14) (20.58)	.574	1.11	20.54	.63	637.82	17.32	18,005
IV	ln SSE	$-38.80\ (\frac{1}{T}) + .63\ \ln ASE + 25.15\ \frac{(1-\rho^2)}{T} - 1.18\ \frac{\sigma^2}{T}$ (2.12) (.033) (5.76) (.54)	.44	2.46	.41	2.86	-.60	4.06	2.43
2S	ln SE	$-.247\ (\frac{1}{T-29}) + .891\ \ln ASE + 36.47\ \frac{\lambda^3\ (1-\rho)(1+\gamma^2)}{T}$ (.042) (.013) (1.52)	.419	2.13	.27	6.85	2.14	8.57	13.83
IGN	ln SE	$-.196\ (\frac{1}{T-29}) + .923\ \ln ASE + 28.42\ \frac{\lambda^3\ (1-\rho)(1+\gamma^2)}{T}$ (.043) (.012) (1.46)	.563	1.18	.26	3.66	1.14	12.41	10.24
IV	ln SSECV	$.43\ (\frac{1}{T}) + .97\ \ln\ ASE$ (.42) (.069)	.99	1.44	.09	143.25*	.99*	*	*
2S	ln SECV	$-.23\ (\frac{1}{T-29}) + .89\ \ln ASE + 36.54\ \frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$ (.043) (.013) (1.55)	.447	2.13	.27	8.79	2.07	8.31	13.83
IGN	ln SECV	$-.19\ (\frac{1}{T-29}) + .92\ \ln ASE + 28.70\ \frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$ (.41) (.013) (1.50)	.576	1.21	.27	5.24	1.04	10.64	10.22

Table 9. Response Functions for λ in Model 2

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(8)$	$Q_{Ai}(8)$	$Q_i(2)$	$Q_{Ai}(2)$
IV	B	$\frac{1}{\sqrt{T}}[\underset{(.11)}{.16}\,\sigma^2 - \underset{(3.45)}{.96}\,\lambda + \underset{(1.19)}{1.71}\,\rho - \underset{(3.91)}{13.90}\,\lambda^2\gamma]$	.20	.74	1.57	3.66	8.80	.02	4.30
2S	B	$\frac{1}{\sqrt{T}}[\underset{(22.56)}{29.99} - \underset{(27.89)}{184.82}\,\lambda^3(1-\rho)(1+\gamma^2)]$	.247	2.39	34.66	3.03	325.10	33.05	16,819
IGN	B	$\frac{1}{\sqrt{T}}[\underset{(19.76)}{52.70} - \underset{(24.44)}{271.47}\,\lambda^3(1-\rho)(1+\gamma^2)]$	.470	1.69	30.37	6.84	176.18	176.18	12,911
IV	BCV	$\frac{1}{\sqrt{T}}[\underset{(.11)}{.48}\,\sigma^2 - \underset{(1.52)}{9.76}\,\lambda - \underset{(.53)}{2.22}\,\rho - \underset{(1.73)}{2.14}\,\lambda^2\gamma]$	.43	1.90	.69	5.02	2.12	.11	1.21
2S	BCV	$\frac{1}{\sqrt{T}}[\underset{(22.60)}{28.10} - \underset{(27.95)}{181.29}\,\lambda^3(1-\rho)(1+\gamma^2)]$	.240	2.40	37.72	4.30	311.85	1.78	16,960
IGN	BCV	$\frac{1}{\sqrt{T}}[\underset{(19.85)}{50.82} - \underset{(24.56)}{267.93}\,\lambda^3(1-\rho)(1+\gamma^2)]$	.461	1.70	30.51	.61	175.82	15.76	12,861
IV	ln SE	$-\underset{(6.19)}{31.32}\left(\frac{1}{T}\right) + \underset{(.033)}{.69}\,\ln ASE - \underset{(.62)}{1.82}\,\frac{\sigma^2}{T} + \underset{(6.54)}{23.72}\,\frac{(1-\varphi^2)}{T}$	.150	2.15	.47	3.50	.26	4.14	.90
2S	ln SE	$-\underset{(.057)}{.255}\left(\frac{1}{T-29}\right) + \underset{(.014)}{.848}\,\ln ASE + \underset{(2.29)}{45.99}\,\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$	.659	2.06	.37	6.23	7.67	1.84	21.43
IGN	ln SE	$-\underset{(.051)}{.140}\left(\frac{1}{T-29}\right) + \underset{(.013)}{.888}\,\ln ASE + \underset{(2.07)}{37.07}\,\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$	.810	1.11	.34	4.01	5.42	3.96	16.88
IV	ln SECV	$-\underset{(.91)}{3.20}\left(\frac{1}{T}\right) + \underset{(.012)}{.93}\,\ln ASE + \underset{(.07)}{.45}\quad(T\leq 30)(\lambda>.7)(\gamma>.7)$	.970	1.66	.19	61.69	1.79	*	*
2S	ln SECV	$-\underset{(.058)}{.27}\left(\frac{1}{T-29}\right) + \underset{(.014)}{.85}\,\ln ASE + \underset{(2.32)}{46.27}\,\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$	.661	2.00	.38	5.68	7.63	46.40	21.43
IGN	ln SECV	$-\underset{(.052)}{.15}\left(\frac{1}{T-29}\right) + \underset{(.013)}{.89}\,\ln ASE + \underset{(2.09)}{37.31}\,\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T}$	.810	1.08	.34	3.58	5.36	3.99	16.88

Table 10. Response Functions for ρ in Model 2

Estimator	Dep. Variable	Response Function	R^2	D.W.	S_i	$Q_i(8)$	$Q_{Ai}(8)$	$Q_i(2)$	$Q_{Ai}(2)$
IV	B	$\frac{1}{\sqrt{T}}[\underset{(8.06)}{6.56} - \underset{(10.26)}{105.51}\rho + \underset{(10.92)}{109.54}\lambda^3(1-\rho)(1+\gamma^2)]$	.740	3.04	11.71	16.06	380.51	9.29	4,623
2S	B	$\frac{1}{\sqrt{T}}[\underset{(4.29)}{-11.88} - \underset{(5.47)}{70.07}\rho + \underset{(5.82)}{53.18}\lambda^3(1-\rho)(1+\gamma^2)]$	.775	2.82	6.24	.716	444.44	7.56	1,433
IGN	B	$\frac{1}{\sqrt{T}}[\underset{(3.87)}{-17.12} - \underset{(4.93)}{71.01}\rho + \underset{(5.25)}{35.08}\lambda^3(1-\rho)(1+\gamma^2)]$	.769	2.28	5.62	2.33	742.34	44.44	912
IV	BCV	$\frac{1}{\sqrt{T}}[\underset{(7.44)}{-7.95} - \underset{(9.46)}{83.70}\rho + \underset{(10.08)}{113.11}\lambda^3(1-\rho)(1+\gamma^2)]$	.738	2.95	10.81	14.20	473.10	6.69	3,996
2S	BCV	$\frac{1}{\sqrt{T}}[\underset{(3.78)}{-12.10} - \underset{(4.81)}{37.02}\rho + \underset{(5.12)}{56.52}\lambda^3(1-\rho)(1+\gamma^2)]$	.713	2.61	5.49	.03	166.71	.36	1,385
IGN	BCV	$\frac{1}{\sqrt{T}}[\underset{(3.39)}{-17.33} - \underset{(4.32)}{37.97}\rho + \underset{(4.60)}{38.42}\lambda^3(1-\rho)(1+\gamma^2)]$	.678	1.86	4.93	10.85	742.34	11.64	912
IV	ln SE	$\underset{(.80)}{-8.30}(\frac{1}{T}) + \underset{(.008)}{.955}\ln ASE + \underset{(1.21)}{24.44}\frac{\rho^2}{T} - \underset{(.39)}{3.10}\frac{\rho}{T}$	.867	1.60	.087	4.13	.11	.24	.039
2S	ln SE	$\underset{(1.09)}{-5.94}(\frac{1}{T}) + \underset{(.011)}{.942}\ln ASE + \underset{(.56)}{6.92}\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T} + \underset{(1.61)}{18.83}\frac{\rho^2}{T}$	.857	1.90	.12	6.61	.39	10.21	1.05
IGN	ln SE	$\underset{(2.02)}{-6.96}(\frac{1}{T}) + \underset{(.020)}{.909}\ln ASE + \underset{(1.03)}{9.07}\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T} + \underset{(2.96)}{16.59}(\frac{\rho^2}{T})$	.655	1.47	.21	8.46	1.56	250.3	15.84
IV	ln SECV	$\underset{(1.45)}{-5.84}(\frac{1}{T}) + \underset{(.015)}{.91}\ln ASE + \underset{(2.17)}{18.76}(\frac{\rho^2}{T}) - \underset{(.50)}{6.02}(\frac{\rho}{T})$	.836	1.73	.11	8,723	.13	2,566	.03
2S	ln SECV	$\underset{(3.48)}{-6.59}(\frac{1}{T}) + \underset{(.035)}{.90}\ln ASE + \underset{(1.78)}{5.79}\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T} + \underset{(5.10)}{20.22}(\frac{\rho^2}{T})$	.340	1.94	.37	1.94	.45	2.42	1.54
IGN	ln SECV	$\underset{(2.29)}{-6.53}(\frac{1}{T}) + \underset{(.023)}{.88}\ln ASE + \underset{(1.17)}{6.62}\frac{\lambda^3(1-\rho)(1+\gamma^2)}{T} + \underset{(3.36)}{19.33}(\frac{\rho^2}{T})$	.543	1.78	.24	10.21	1.65	194	16.01

Fig. 2. The logarithm of the Estimated Standard Errors (ESE) for the 2S estimator of λ (Model 2). (For odd numbers, T = 30 , for even numbers, T = 60).