The Korean Economic Review, Vol. 4, No. 1

STOCHASTIC REDUCED FORM FORECASTS AND THEIR RISK IMPROVEMENTS UNDER STRUCTURAL SPECIFICATION UNCERTAINTY IN A SIMULTANEOUS EQUATIONS MODEL

JIN-HO JEONG*

ABSTRACT

In the context of a linear simultaneous equations model, this study indetifies two major reasons for the generally poor predictive performance of almost all the traditional reduced form forecasts. Some of them do not have finite moments in small samples and hence adversely affected from the departure from normality. All of them treat overidentification restrictions in an exact manner, and hence produces erratic behavior under misspecification. Forecasts based on these reduced form estimators are also subject to unbounded risk in these situations since they are the linear combination of reduced form estimators. In search of reduced form forecasts with lower risk consequences, this study compares the parametric and nonparametric risk measures of the forecasts in Monte Carlo experiments. The forecasts include the recently proposed the Modified Stein-like Reduced Form (MSRF) and the Generic Reduced Form (GRF) estimators, along with unrestriceted, restricted, and partially restricted reduced form forecasts. And the experiments are controlled by the key model parameters developed in the small sample theory. The results of this study show that the lower centrality and/or misspecification, rather than the sample size change that used frequently in the asymptotic theroy, is the main source of the unbounded risks of traditional reduced form estimators. Especially the traditional 2SLS and 3SLS based forcasts have non-negligible probability for producing outliers. To summarize the result of this experiment, the lower risk reduced form forecasts, which have flexibility in adopting uncertain structural information according to its sample validity, should be used for prediction or policy analysis.

I. INTRODUCTION

In the context of a linear simultaneous equations model, this study examines two major reasons for the generally poor predictive performance of almost all the traditional 'restricted (or Derived)' Reduced Form (DRF) estimators which are derived from structural coefficient estimates. First, in finite samples, the non-existence of moments for some estimators results in unbounded forecast intervals. Second, serious misspecifications affect forecasting adversely in a number of ways. Unlike these

* Associate professor at Economics Department of Economics, University of Cincinnati.

derived reduced form estimators, the 'Uurestricted' Reduced Form (UFR) estimator (which is computationally simple but inefficient) possesses moments of all orders and is consistent even under certain types of misspecifications.

Recently two lower risk reduced form estimators for a linear simultaneous equation model have been proposed by Maasoumi (1978, 1983, 1986). One is the Modified Stein-like Reduced Form (MSRF) estimator which combines URF and DRF estimators according to a Wald-type specifcation test. The other is the Generic Reduced Form (GRF) estimator which mixes URF and DRF via a matrix weight of precision contributed by sample and non-sample information. The statistical improvement of these estomators in terms of general quadratic loss criteria is obtained through finite sample justification and explicit control over specification uncertainty. It is also expected that forecasts based on these lower risk reduced form estimators would outperform, in terms of quadratic risk, forecasts based on various standard reduced form estimators. Maasoumi and Jeong (1988) reports some properties of these and other traditional structural form and reduced form estimators in macroeconometric applications and in experimental studies.

An experimental investigation of the predictive performance of these combined (mixed) estimators along with other unrestricted, restricted (derived), and partially restricted reduced form estimators is the prime concern of this study. Analytic results from the small sample theory are used carefully to guide the expermental design in order to minimize the loss of generality faced in usual Monte Carlo studies. Sargan (1976a), Anderson, Morimune and Sawa (1983), Mariano (1982) and phillips (1983) uncovered the key model parameters from the exact probability density function of a single equation instrumental variables estimator, which affect the location, dispersion, and shape of the density. Rhodes and Westbrook (1981), Maasoumi and Phillips (1982) and Knight (1986) examined the effect of exclusion type misspecifcation on the density through changes in these parameters. The different experimental situations examined here are characterized by the changes in these parameter values from those under standard assumptions such that experiments become more controllable.

The predictive distributions obtained from the standard and nonstandard reduced form coefficient estimates are examined under three types of situations which occur frequently in applied econometrics : (1) small sample size hurting asymptotic justification, (2) lower concentration causing departure from normality, and (3) exclusion-inclusion type misspecifications. The experimental results under these conditions are compared relative to a standard situation using parametric and nonparametric statistics, and some of the important results are highlighted by contrasting the sampling densities and cumulative distributions of the lower risk reduced form forecasts to those of the traditional reduced form forceasts.

Major conclusion drawn from these experiments is that the nonstandard combined reduced form forecasts outperform the traditional restricted, partially restricted, and unrestricted reduced form forecasts under general risk rankings, except some tied rankings of the GRF forecasts under *unsure* specifcation with the DRF-FIML forecasts under reduced sample size and under lower centrality. Risk rankings of various forecast distributions are presented using both parametric and nonparametric criteria : the root mean squared error (the standard deviation of forecasts) ranking and the quintile (80% range length) ranking . Contrarily to traditional beliefs in analytic studies, the FIML forecast performs very well in small samples under correct specification, and reasonably well under low centrality and even under misspecifcation. The reasonable performance of FIML predictors may be due to existence of finite moments and the higher centrality, because the DRF-FIML forecast is the only standard efficient SEM forecast which has some finite moments. The DRF-2SLS forecast exhibit the worst performance in this experimental study. This is an important result when we consider its popularity amongst applied macroeconomists despite more than a decade warnings about its unbounded forecast interval from small sample theoriests.

The paper is organized as follows. After a brief discussion about forecasts, various types of reduced form estimators are introduced along with those characteristics relevant to the distribution of forecasts in the next section. The experimental design is summarized in Section 3. And discussions on the experimental result and conclusion follow in the sections 4 and 5.

Ⅱ. FORECASTS AND REDUCED FORM ESTIMATOR

The aim of predictive inference is believed not to fit a probability model to the data, but rather to choose a probability distribution which depicts well the random behavior of future samples that may be either independent of or correlated with the observed data. This view is shared by Aitchison (1975), Akaike (1977) and Larimore (1983) among others. In the context of forecasting from a simultaneous equation model it is well known that 'good fit' in terms of structural residuals does not necessarily imporve the prediction of a random phenomenon different from but related to that being observed. The fundamental problem is that the choice of the reduced form estimator gives different probability distributions of forecasts for the same specifications and data employed. The implication is that the same nonsample and sample information may not have the same predictive inference. Suppose that a vector of n endogenous variables, y_t, $t=1,2,\ldots,T$, is generated by the reduced form equation $y_t = P'z_t + v_t$, where z_t is a vector of m non-stochastic exogenous variables with $Z'Z/T = M + O(1/T)$, M being positive definite, and v_t is a set of random normal errors, that are serially independent, and distributed $N(0,\Omega)$ for each t. P is an $m \times n$ matrix of reduced form coeffcients. The 'h-period ahead forecast' of y based on available information at time T, $\hat{y}_{T+h}$ is given by

$$(1) \qquad \hat{y}T + h = \hat{P}'z_{T+h}$$

where z_{T+h} is a vector of a priori known values of m exogenous variables for time $T+h$ and $\hat{P}$ is any consistent estimator for P. The forecast error for a one-period ahead forecast is given by

$$(2) \qquad \hat{v}_{T+1} = \hat{y}_{T+1} - y_{T+1} = (\hat{P}-P)'z_{T+1} - v_{T+1}$$

The forecast error is therefore decomposed into two parts : the first part representing the errors in estimating P by a consistent procedure which will be discussed below, and the second part representing the disturbances occuring, at time $T+1$. Vectorizing (2) with respect to P gives

$$(3) \qquad \hat{v}_{T+1} = (1/\sqrt{T})(I_n \otimes z'_{T+1})\sqrt{T}(\hat{p}-p) - v_{T+1}$$

where p denotes a mn dimensional column vector stacking the columns of P one underneath the other to form a single vector and $\otimes$ denotes the Kronecker product. For a chosen consistent and asymptotically unbiased estimator $\hat{P}$ for P, as shown by Goldberger, Nagar and Odeh (1961) the forecast error vector has zero asymptotic mean,

$$(4) \qquad E(\hat{v}_{T+1}) = 0,$$

with asymptotic variance covariance (AV) matrix

$$(5) \qquad AV(\hat{v}_{T+1}) = \frac{1}{T}(I_n \otimes z'_{T+1})AV(\sqrt{T}(\hat{p}-p))(I_n \otimes z'_{T+1}) + \Omega.$$

Using these results, we know that the efficiency of the forecast is related to the efficiency of the reduced form coefficient estimator for a given set of future values of exogenous variables z_{T+1}. Not only the asymptotic properties but also the small sample properties of the reduced form estimators directly affect the properties of the forecasts. Therefore we will examine the various types of reduced form coefficient estimators in order to make conjectures about the probability distributions of the forecasts.

1. Unrestricted, Restricted and Partially Restricted Reduced form Forecasts

Unrestricted Reduced Form Forecast.

If P consists entirely of unknown constants and does not depend on any other parameters, the reduced form model becomes a special case of seemingly unrelated regression model of Zellner (1962) with identical regressors. It is well known that the ordinary least squares estimator applied to this model is as efficient as Aitken's generalized least squares estimator. The unrestricted least squares (ULS) reduced form estimator of P is given by

$$(6) \qquad \hat{P}_{ULS} = (Z'Z)^{-1}Z'Y$$

with the asymptotic distribution

$$(7) \qquad \sqrt{T}vec(\hat{P}_{ULS}-P) \sim N(0, \Omega \otimes M^{-1})$$

where $X=[Y\ Z]$ denotes $T \times (n+m)$ data matrix, and $plim(Z'Z/T)=M$, M being a positive definite constant matrix. The ULS estimator ignores individual structural differences of the endogenous variables but makes its forecast error orthognal to the design matrix of all exogenous variables and exhibit minimum variance property for whatever data given. As long as data has full rank, its forecast also has finite moments of all orders. It is inefficient but computationally simple and consistent if the design matrix contains all relevant exogenous variable.

Restricted Reduced Form Forecast.

If P consists of unknown constants but depends on other elements of the included (unconstrained) structural parameter vector α from the complete structural form, then the reduced form coefficient estimates may be derived from the consistent and efficient structural estimates of α. The restricted (derived) reduced form (DRF) estimator of P is given by

$$(8) \quad \tilde{P} = P(\tilde{\alpha}) = -\tilde{C}\tilde{B}^{-1}$$

where the complete structural form, $YB + ZC = XA = U$, with $E[U] = 0$, and $E[U'U] = \Sigma$, is assumed to have the $(m+n) \times n$ structural coeffcient matrix $\tilde{A} = [\tilde{B}' : \tilde{C}']$ consisting of zeros and the elements of particular estimates of α identifiable from the reduced form coeffcient parameters.

The structural form representation of the sample information X, $XA = U$, explicitly models the joint endogeneity that exists among elements of y_t. Under the assumption that no restriction is imposed on covariance matrices, following Sargan (1976a, P. 437), we may reparametrize the overidentifying restrictions as the "specification projection" from the included sturctural coeffcient vector α into the restricted structural coeffcient matrix A:

$$(9) \quad vec(A) = s - S\boldsymbol{\alpha}$$

where s normalizes the diagonal elements of B, and S imposes linear dependency, or simply exclusion restrictions. For the latter case S has dimension $n(n+m) \times (n(n+m)-q)$, where q represents the number of linearly independent restrictions. We call (9) as 'specification constraints' since it summarizes the structural link of the jointly endogenous variables.

In terms of this specification constraints, we now consider the 3SLS coeffcient estimator of α

$$(10) \quad \tilde{\boldsymbol{\alpha}}_{3SLS} = (S'(\hat{\Sigma}^{-1} \otimes \hat{X}'X)S)^{-1}(S'(\hat{\Sigma}^{-1} \otimes \hat{X}')y)$$

which has asymptotic distribution

$$(11) \quad \sqrt{T}(\tilde{\boldsymbol{\alpha}}_{3SLS} - \boldsymbol{\alpha}) \sim N(0, (S'(\hat{\Sigma}^{-1} \otimes Q'MQ)S)^{-1})$$

where $\hat{X} = Z[\hat{P} : I] = Z\hat{Q}$, $\hat{Q} = [\hat{P} : I]$, is the ULS projection of Z into X with $plim\ \hat{Q} = Q$, and $\hat{\Sigma}$ is any consistent estimator of Σ.

Now we substitute the specification constraints $\mathrm{vec}(\tilde{A} - A) = S(\tilde{\alpha} - \alpha)$ into the vectorized form of the identity $\tilde{P} - P = -\tilde{Q}(\tilde{A} - A)\tilde{B}^{-1}$, as in Dhrymes (1973, p121). Then we obtain the following relations.

(12) $\sqrt{T}\,(\tilde{p}_{3SLS} - p) = -(\tilde{B}'^{-1} \otimes \tilde{Q})\, S \sqrt{T}\,(\tilde{\alpha}_{3SLS} - \alpha)$

from which the asymptotic distribution of the derived reduced form based on 3SLS (DRF—3SLS) can be obtained.

Asymptotically, the derived reduced form estimator $\tilde{P}$ behaves like a linear transformation of the structural estimator $\tilde{\alpha}$ as explained in Rao (1973) in the another context.

So the efficiency of the reduced form coefficient estimator, and hence that of the predictor, is directly related to the efficiency of the structural coefficient estimator. When the information contained in the contemporaneous covariance matrix $\Sigma = B'\theta B$ is ignored and $\hat{\Sigma}$ in (10) is replaced by an identity matrix I, the resulting 2SLS estimator (DRF-2SLS) becomes asymptotically less efficient. When the ULS projection $\hat{X} = Z[\hat{P} : I]$ in the same equation is replaced by the iterative DRF projection $\tilde{X} = Z[-\tilde{C}\tilde{B}^{-1} : I]$ based on the jointly estimated structural FIML estimates $\tilde{\alpha}_{FIML}$, the resulting FIML reduced form estimator (DRF-FIML) becomes more efficient as recognized by Hausman (1975).

However, these asymptotic properties of the DRF estimators usually break down in small samples. If there are overidentifying restrictions on the structural equations, the derived reduced form coefficients from 2SLS or 3SLS will in general possess no integer moments (Sargan; 1976b). The fact that the FIML estimator is independent of the normalization rule means that the FIML estimate of a structural coefficient can be interpreted as the reciprocal of another FIML estimate under a different normalization. This implies that the distribution of such an estimator has no integer moments. Mariano (1982) and Phillips (1983) provide an excellent guide on the non-existence of moments issues and other finite sample results. These properties influence the probability of outliers in the derived reduced form coefficients and hence directly affect the distribution of the forecasts adversely, because the forecast interval depends, among other things, on the first two moments (mean and variance) of the reduced form coefficient estimators.

Partially Restricted Reduced Form Forecast.

If P incorporates overidentifying restrictions of a single structural equation at a time, then the partially restricted reduced form (PRRF), based on $(I \otimes Q)(s - S\alpha) = 0$, or $P = (I \otimes Q)S\alpha$, would be an appropriate estimator. As an approximation, we substitute $\hat{Q} = [\hat{P} : I]$ for Q and α_{2SLS} for α, then we get the PRRF estimator p_{PRRF} of p

(13) $\tilde{p}_{PRRF} = (I \otimes \hat{Q})\, S\, \tilde{\alpha}_{2SLS}$

which has $\hat{P}_{i,\,PRRF} = \hat{Q}\, S_i\, \hat{\alpha}_{i,\,2SLS}$ of Kakwani and Court (1972) as a special case when specification constraints are separable.

One feature of PRRF is that it can avoid serious misspecification by replacing a doubtful structural equation with an unrestricted reduced form equation. However Kakwani (1975) and Knight (1983) showed that PRRF is neither efficient nor inefficient relative to ULS and is less efficient relative to DRF-2SLS. S-B. Park (1982) studies a forecasting property of PRRF that it produces the same forecast as ULS and DRF-2SLS when exogenous variables have the values of their sample mean in the forecast period.

2. Improved Reduced form Forecasts

The fact that many of the popular DRF estimator possess no integral moments and that standard efficient estimators are irresponsive to structural misspecification has led to the search for the lower risk estimators. Members of the class of lower risk reduced form estimators in Jeong (1985) are the reduced form coefficient estimators which possess finite moments up to some orders in small samples and which, by their construction, incorporate a priori structural information according to either data admissibility for given specified significance levels or the assessed level of specification uncertainty. This class has the mixed (or combined) reduced form, a mixture of the URF and the DRF, as its typical form. These estimators combine ULS and DRF estimators in various ways in the hope that the behavior of the predictive distribution of these combined estimator is improved.

Pretest Based Reduced Form Forecast.

If P is subject to structural misspecification, then the use of the Modifed Stein-like Reduced Form (MSRF) estimator of Maasoumi (1978) will combine the ULS and DRF-3SLS estimators, by a shrinkage parameter λ such that

$$(14) \qquad \hat{P}_{MSRF} = \lambda \hat{P}_{3SLS} + (1 - \lambda) \hat{P}_{ULS}$$

where

$$\lambda = \begin{cases} 1, & \text{if } \phi \le C_p \\ \sqrt{C_p / \phi}, & \text{if } \phi > C_p. \end{cases}$$

The scalar weight depends on the preliminary test of the overidentifying restrictions using Wald-type specfcation test statistics

$$(15) \qquad \phi = tr\,(\hat{\Omega}^{-1}(P-P_{3SLS})'(Z'Z)(P_{P3SLS}))$$

where $\hat{\Omega} = Y'(I - Z(Z'Z)^{-1}Z')Y/T$ and C_p is the asymptotic critical value of the test corresponding to the chosen significance level. Under serious structural misspecification the value of λ will become smaller than 1 and the MSRF forecast will be pulled toward the ULS forecast, and hence it is deemed necessary to discard some portion of the structural specification in favor of sample information.

The MSRF estimator has a simple struture and its implication is unambiguous. The scalar stochastic mixing weight combines URF and DRF only in reference to the validity of the parameter constraints for a specified significance level. As demomstrated in Maasoumi (1978), it possesses the first $(T - n - m)$ moments in finite samples, that is, as many moments as the DRF-FIML estimator. Moreover, its limiting moments approximate the moments of the asymptotic distribution of DRF-3SLS with reasonable sample size. Hower, it is still interesting to examine the sensitivity of MSRF to the erratic small sample behavior of the Wald type asymptotic test on which MSRF is based.

Generic Reduced Form Forecast

If P is specified as being generated by a practically known distribution function so that the vectorized coefficient constraint holds only stochastically, that is, $E[PB+C]=0$ or $\bar{P}B+C=0$, $\bar{P}$ being the mean of P, then the GRF estimator of p provides the predictor of the matrix weighted average form which is well known in the Bayesian analysis of the multivariate regression model (Chamberlain and Leamer (1976), for example). Maasoumi's (1983, 1986) (empirical) choice of the GRF estimator of p is given by

$$(16) \quad \hat{p}_{GRF} = W^{*}\hat{p}_{ULS} + (I - W^{*})\hat{p}_{3SLS}$$

where W^{*} *is the* $nm \times nm$ stochastic weight matrix which reflects the precision contributed by the sample information in relation to the posterior precision matrix contributed by prior specification precision Ω_d^{-1}, which in turn depends inversely on c, the experimentally determined degree of specification uncertainty, and proportionally on the square root of the sample size. See Jeong (1985) or Maasoumi (1986) for details for the derivation of W^{*}. In our experiments we choose $c=1.0$ for 'unsure' specification and $c=0.05$ for 'certain' specification.

Under uncertain structural specification, the forecast based on $\hat{P}_{GRF}$ may reduce to the ULS forecast as $c\rightarrow\infty$ or to the DRF-3SLS forecast as $c\rightarrow 0$. The limiting properties of $\hat{P}_{GRF}$ indicate these features. For given initial choice of Ω_d which has smallness of order $1/\sqrt{T}$ or less ($T^{-3/2}$ in our experiment), W^{*} converges to I such that $\hat{P}_{GRF} \rightarrow \hat{P}_{ULS}$ as sample size increases. For given T, W^{*} converges to 0 such that $\hat{P}_{GRF} \rightarrow \hat{P}_{3SLS}$ as specification uncertainty is reduced ($\Omega_d\rightarrow 0$), and W^{*} converges to I such that $\hat{P}_{GRF} \rightarrow \hat{P}_{ULS}$, as specification uncertainty increases ($\Omega_d\rightarrow\infty$). Given that $\operatorname{limit}_{T\rightarrow\infty} T\Omega_d$ is a constant, $\bar{W}^{*}$ converges to some limit, W^{*}, such that a mix of $\hat{P}_{ULS}$ and $\hat{P}_{3SLS}$ defines $\hat{P}_{GRF}$. Its asymptotic variance matrix, AV, is given by

$$(17) \quad AV(\hat{p}_{GRF}) = AV(\hat{p}_{3SLS}) + \bar{W}^{*}[AV(\hat{p}_{ULS}) - AV(\hat{p}_{3SLS})]\,\bar{W}^{*}.$$

In finite samples the GRF estimator $\hat{P}_{GRF}$ may have finite integer moments up to order $k/2$, k being the total degrees of overidentification, (as many moments as the structural 3SLS estimators can possess under the same situation). By its design, the GRF forecast is responsive to structural misspecification, so it is expected to behave well in small samples.

Differences in the modelling environments and in the design of the reduced form estimators will directly affect the predictive distributions. The choice of the reduced form estimator matters in forecasting, i.e., the formation of future expectations. We want to examine the forecasting properties of these estimators in the experimental situation which resembles key features of the reality.

Ⅲ. EXPERIMENTAL DESIGN, SMALL SAMPLE THEORY, AND MONTE CARLO METHODS

Different experimental designs and Monte Carlo methods may favor particular estimators, hence the conclusions from these experiments may be 'design' specific and are not as informative as they could be for general use. Therefore, in experimental design, it is important to reduce the parameter spaces to an essential set in order to identify the critical parametric functions which influence the shape of the relevant small sample distribution. The usual procedure in analytic small sample theory is to implement the standardizing transformation which reduces the sample second moment matrix to the sample proportional identity matrix and the covariance matrix of the disturbance vector to the canonical form. In our experimental design, the conditions under which predictors perform differently are characterized by the five key model parameters which are frequently used to represent the canonical form of the exact distributions of simpler estimators. We consider the following five key model parameters to describe the changes in our experimental situations : (a) the sample size, (b) the sampl cross moment matrix of exogenous variables, (c) the degrees of overidentification, (d) the non-centrality (concentration) parameter[1] and (e) the sample variability of the systematic components of Y relative to random components. We introduce two basic models whose predictors from eight reduced form estimators are examined in four different experimental situations.

[Table 1] Two Basic Experimental Models

	Model A	Model B
(a) structural parameters		
$a' = [a_1' : a_2']$	$[b_1 c_1 c_2 : b_2 c_7 c_8]$	$[b_1 c_1 c_2 c_3 : b_2 c_8]$
	$= [-.5\ \ .5\ \ -.75\ \ -4.0\ \ 4.0\ \ -1.6]$	$= [-.5\ \ .5\ \ -.75\ \ 4.0\ \ -4.0\ \ -1.6]$
$A' = [B' : C']$	1 -.5 : .5 -.75 0 0	1 -.5 : .5 -.75 4.0 0
	-4.0 1 : 0 0 4.0 -1.6	-4.0 1 : 0 0 0 -1.6
Σ	1.0 .5	same as Model A
	.5 1.0	
(b) reduced form parameters		
P'	.5 -.75 2.0 -.8	.5 -.75 4.0 -.8
	2.0 -3.0 4.0 -1.6	2.0 -3.0 16.0 -1.6
Ω, Ω-1	1.75 6.0 28.0 -8.0	same as Model A
	6.0 21.0 -8.0 2.33	
(c) log-likelihood values		
2 $ln\|B\| - ln\|\Omega\|$	(2.0) (0.0) $- ln$ (.75) $= 0.2877$	same as Model B

1) We call $P'(Z'Z/T)P$ a noncentrality parameter matrix. However the scalar 'noncentrality parameter', or 'concentration parameter', μ^2, in the small sample theory literature is defned as

$$\mu^2 = TP'_{22}P_{22}$$

Table 1 shows two basic experimental models which have two endogenous variables and four purely non-stochastic and orthogonal exogenous variables with zero sample mean. These models are linear and share identical disturbance structures. The distribution of the first endogenous variable is less disperse than that of the second. The reduced form coefficients of the third exogenous variable are changed due to identifying restriction swaping that variable from the first equation to the second. In the frst model (Model A), the structural form of both equations are overidentified by degree one, and, in the second model (Model B), the first structural equation is just identified and the second is overidentified by two degrees. Logarithm of the likelihood value for chosen true parameter values are same for both models such that distributions of different forecasts are directly comparable.

Table 2 shows the experimental values of the key model parameters for the basic model (Experiment Ⅰ) and their changes on coefficient experiments. The first (Experiment, Ⅰ) is an experiment under correct specificaton and medium sample size of 48. This is the most favorable situation for the traditional estimators which rely heavily on asymptotic theory and exactness of coefficient constraints. The second (Experiment Ⅱ) is an experiment under correct specification but reduced sample size of 24. The change in sample size will affect the sample moment matrix of Z, however it does not affect the concentration parameter in our design. The third (Experiment Ⅲ) is an experiment under correct specification, in small samples and with lower centrality. By scaling down the exogenous variables, the determinant of the sample second moment matrix, and hence that of the concentration parameter matrix will become small, so this lower central tendency cause the departure from normality. The fourth (Experiment Ⅳ) is an experiment where structural misspecification is introduced by switching models. Hence, when model A is perceived as model B, the first equation becomes overidentified by one degree less while the second equation by one degree more. In this situation, wrongful inclusion of exogenous variable causes the first equation to be just-identified and wrongfiul exclusion causes the second equation to be over-identified by two degrees. Under structural misspecification and in small samples, the behavior of the predictive distributions becomes unpredictable because of the presence of frequent outliers in the reduced form estimators on which our forecasts are based.

The conditional forecasts are generated from the estimated reduced form coefficients with a priori known exogenous variables. The behavior of the individual reduced form coefficient estimator and that of the forecast would be different even in static models, in the sense that the the latter depends on all members of the former. Therefore the behavior of the reduced form coefficients examined in Jeong (1985, Chapter Ⅳ, its summary is also reported in Maasoumi and Jeong (1988)) would guide our conjectures about the behavior of forecasts only in a limited sense.

---

where $Z'Z=TI$ is assumed and P_{22} denotes the reduced form coefficient matrix of included endogenous variables on the excluded exogenous variables of the equation under consideration. The distribution of the instrumental variables estimator concentrate more as $\mu^2 \rightarrow \infty$, even if the sample size T remain fixed. See Basemann (1963). Since we consider both IV and ML estimators and their mixtures, we consider $P'(Z'Z/T)P$ as a whole.

[Table 2] The Change of Model Parameters in Four Experiments

Key parameters	Model A		Model B	
Experiment Ⅱ (small samples, higher centrality, correct specification)				
(a) sample size, T	$T=24$		same as Model A	
(b) sample second moment matrix, $\|Z'Z\|$	$\|Z'Z\|=\|24I_4\|=24^4$		same as Model A	
(c) degree of overidentificarion, $v=v_1+v_2$	$2=1+1$		$2=0+2$	
(d) noncentrality parameter,	5. 45	12. 53	17. 45	68. 54
$P'[Z'Z/T]P$	12. 53	31. 56	68. 54	271. 59
(e) sample variability,	52. 44	98. 37	-59. 58	-253. 67
Ω-1 $(P'Z'ZP)/T$	-14. 39	-26. 60	20. 29	85. 41
$tr(\Omega^{-1}(P'Z'ZP)/T$	25. 83		25. 83	
Experiment Ⅰ (medium samples, higher centrality, correct specification)				
(a) T	$T=48$		same as Model A	
(b) $\|Z'Z\|$	$\|Z'Z\|=\|48I_4\|=48^4$		same as Model A	
Experiment Ⅲ (small samples, lower centrality, correct specification)				
(b) $\|Z'Z\|$	$\|Z'Z\|=\|I_4\|=1.0$		same as Model A	
(d) $P'Z'ZP/T$	. 22	. 52	. 73	2. 86
	. 52	1. 32	2. 86	11. 30
(e) Ω-1 $(P'Z'ZP)/T$	2. 19	4. 10	-2. 48	-10. 57
$tr(\Omega^{-1}(P'Z'ZP)/T$	-. 60	-1. 11	. 85	3. 56
	1. 076		1. 076	
Experiment Ⅳ (small samples, higher centrality, misspecification)				
perceived structural model	Model B		Model A	
(c) $v=v_1+v_2$	$2=0+2$		$2=1+1$	

In our experiment, we consider one-period ahead conditional forecasts. The future values of non-stochastic exogenous vaiables are given by

$$z'_{T+1} = [\ 1 \quad 1 \quad \sqrt{2} \quad 0\ ].$$

Hence the true values of y at T+1 for model A and B are

$$y'_{T+1,A} = [\ 2.578 \quad 4.657\ ]$$
$$y'_{T+1,B} = [\ 5.407 \quad 21.629\ ]$$

For the experiment with lower centrality, we divide these by 24.

For eight SEM reduced form estimation techniques, two models in four experiments, eight models in total, are estimated, forecasted and simulated on CDC 170/855 under NOS operating system. I used the Fortran program called the LRRF (Lower Risk Reduced Form) program that I developed for the non-standard SEM estimators such as PRRF, MSRF, and GRF adopting some of the subroutines for the standard SEM estimators from Wymer's SIMUL program.[2] Even though we used the program for the two equation model for experimental purpose, Jeong (1988) demonstrates its use for macroeconometric modelling under uncertain structural specification. For any reasonable size of the standard linear simultaneous equations model, the LRRF program calculates the structural form and reduced form coefficient estimates and their variance matrices, and related statistics, residuals and fitted values, and one period lead forecasts for the following eight estimation techniques : (1) the unrestricted least squares reduced form estimation, (2) the ordinary least squares estimation, (3) the two stage least squares estimation, (4) the partially restricted reduced form estimation, (5) the three stage least squares estimation, (6) the modified Stein-like reduced form estimation, (7) the generic reduced form estimation for specified degrees of specification uncertainty, and (8) the full information maximum likelihood estimation.

In conducting simulation, multivariate normal random deviates with zero mean vector and specified positive definite covariance matrix are obtained from the IMSL subroutine GGNSM. The starting seed which initiate the random number generation are exactly same across all experiments, hence the generated endogenous variables (through $Y = ZP + V$) remain fixed in each of eight estimation techinques. As a result all sampling distributions of forecasts are comparable to each other. In obtaining the empirical cumulative distribution function, two points were examined carefully. Firstly, the chosen random deviates should not deviate from its specfied multivariate normal distribution such that any deviation would attributable to the behavior of reduced form estimators not to random numbers. Secondly, the number of replications should be sufficient to stabilize the tail area behavior, since the erratic behavior of forecasts has nonzero probability of producing outliers. The frequency count on a grid plain and its contour map of the bi-variate normal random deviates indicates that 250 replications chosen for this study is sufficiently informative to examine the sampling properties of various forecasts. As a result the FIML iterative procedure with a modified version of the Newton-Raphson algorithm in the LRRF program are all successful for 250 replications in all eight experiments.

2) The LRRF program was developed in 1984 while I was conducting my dissertation research on the improved reduced form estimators under Professor Esfandiar Maasoumi at Indiana University, Bloomington. Special thanks goes to him for guiding me to this issue. For the experiments specified in this study, the LRRF program estimates and predicts with eight different techniques for 250 replications in about 83.5 CPU running seconds for each experiment.

Ⅳ. THE EXPERIMENTAL RESULTS

The eight sampling experiments, each of which consisting of 250 replications of forecasts for two endogenous variables for the eight estimation techniques, generate substantial amount of computer output. However, they examine only a series of eight sampled set of parameter values from the continuum of parameter spaces. How can we obtain more generally acceptable results from these limited experiments? Experiments conducted here resemble the modelling situation which are encountered frequently in empirical research, e. g., small samples, departure from normality, and misspecification, except for the consideration of dynamism which is left for the future study. To identify the sources of erratic behavior at the tail area in unbonded forecasts, the key model parameters and its link to reduced form estimators are considered. Also the computer program developed for this study enable us to use identical sequences of the generated multivariate normal distributions across experiments, and hence it makes the experimental results comparable across experiments and among forecasts. By this construction, we tried to achieve more general informative conjectures about the choice of predictive distributions through the use of limited but controlled set of Monte Carlo simulations.

Experimental results are reported in two forms : Tables of descriptive and range statistics and graphs of cumulative relative frequencies. For each experiment, tables report statistics for the two endogenous variables in model A and model B. The predictive distributions are described in terms of the statistics of the first four central sample moments, median, and nonparametric range statistics. The deviation of the sample mean from the true known values of the forecast will give the bias of the forecasts generated from alternative reduced form estimators. The standard deviation is nothing but the root mean squared error (RMSE) of the forecasts, which has been a popular quadratic risk measure.

It is clear that for some forecasts, namely those whose first two moments are unbounded, the use of this parametric measure is invalid since its decomposition into variance and squared bias depends on the existence of finite moments. However, the numerical values in sampling experiment are finite. So we report the descriptive statistics of these unbounded forecasts, except for the worst forecasts which are based on the derived reduced form of the structural OLS estimates. The skewness and the (central) kurtosis measures may suggest some departures from normality when they differ from zero. The comparison of the inter-range interval indicates the thickness of tail areas to provide more detailed knowledge of the predictive distribution.

Figures present the cumulative relative frequencies of the forecasts from the standard reduced form estimators (ULS, DRF-2SLS, DRF-3SLS, and DRF-FIML) and the nonstandard reduced form estimators (PRRF, MSRF, GRF ($c=1.0$) and GRF ($c=0.05$)), as well as their relative frequencies for Experiment Ⅲ. The cumulative relative frequencies approximate the cumulative predictive distributions upon which our choice of predictor is based. The tail behavior represents the probability of producing outliers. In general the thicker tails at both edges and the lower slope will indicate frequent outliers and large dispersion. The forecast corresponding to the cumulative relative frequency value of 0.5 will give the median value of the forecast which is usually close to the true value or the sample mean for consistent estimators under normality. The relative frequency figures at 25 sample mid-points are obtained from the cumulative distributions, and give some idea about location, dispersion, and shape of the p.d.f. of the forecasts. Especially the probability of producing outliers beyond the range shown on the figure is plotted as an 'outlier relative frequency' at the both end of the empirical p.d.f.

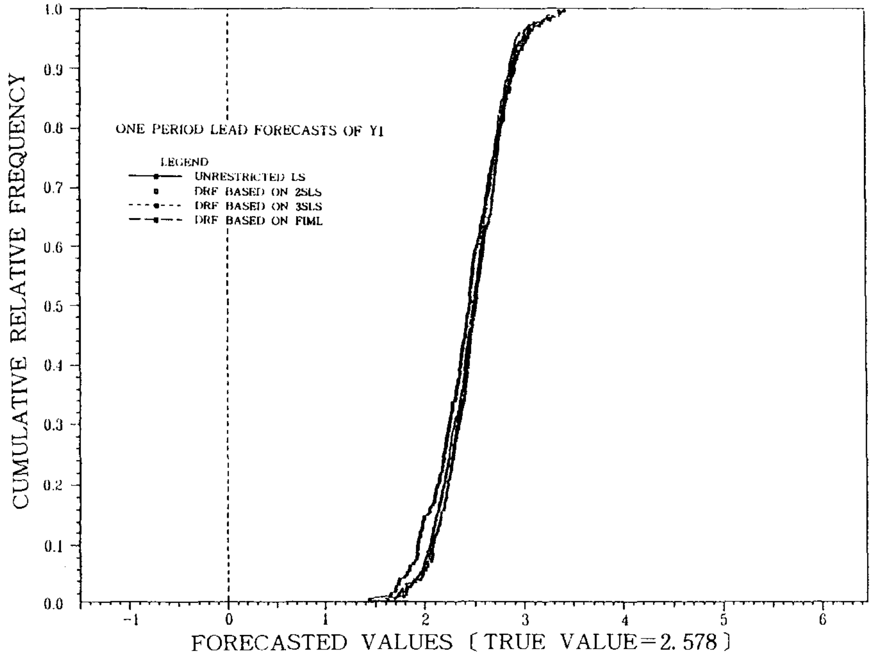
Tables 3 through 6 are tabulated for Experiments I through Ⅳ correspondingly. Experiment I follows standard assumptions with medium sample size ($T=48$), higher centrality (due to $Z'Z=48I_4$), and correct specification. Experiment Ⅱ has a smaller sample size ($T=24$) compared to Experiment I, while Experiment Ⅲ has lower centrality (due to $Z'Z=I_4$) compared to Experiment Ⅱ. And Experiment Ⅳ is misspecified by switching the perceived models, compared to Experiment Ⅳ is misspecifed by switching the perceived models, compared to Experiment Ⅱ. From Experiment Ⅱ to Experiment Ⅳ, several factors cause the model parameters to have widely different values from those of Experiment I as summarized in Table 2. In order to examine the behavior of alternative predictors under different situations, Experiment Ⅱ is compared with Experiment I for the effect of reduced sample size, Ⅲ with Ⅱ for departure from normality, and Ⅳ with Ⅱ for misspecification. In order to save the space, figures are reported selectively. Experimental results are discussed first and suggestions on the choice of predictor are left to the conclusion section.

1. Distributions of Forecasts Under Standard Assumptions

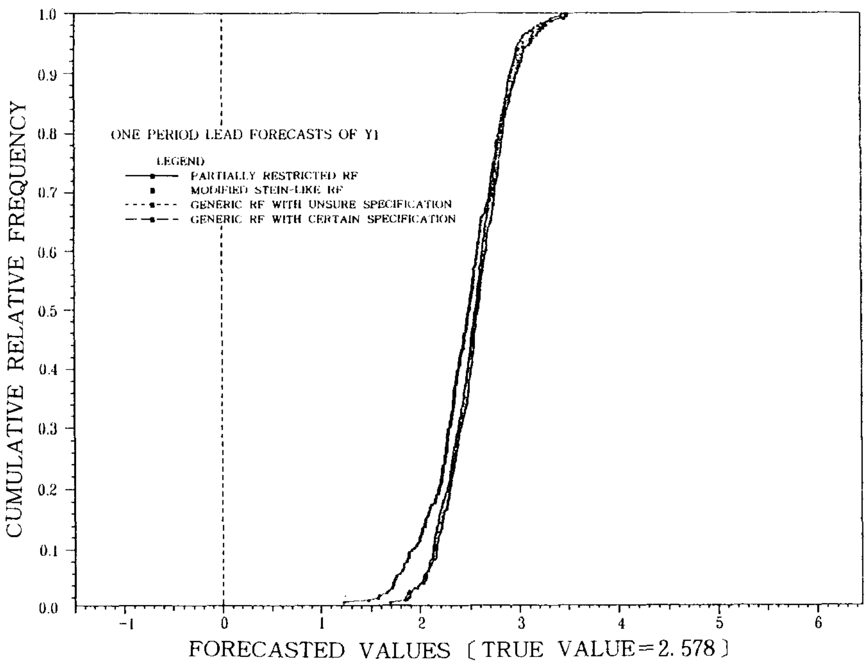
Table 1 and Figures 1 through 4 (for model A) report the results from Experiment I. Under correct specification and with medium size samples ($T=48$), all estimators are expected to exhibit their best performance. The asymptotic equivalence of FIML and 3SLS, and hence the equivalence of their reduced form forecasts is clear in terms of descriptive statistics from Table 1. Though the forecasts based on 3SLS are less biased than those based on FIML, both have very close median values, while the FIML forecasts are closer to a normal distribution in terms of skewness and kurtosis. Under correct specification, the validity of the overidentifying restrictions cannot be rejected in MSRF, and hence, with the value of λ being equal to one, the MSRF forecasts become identical to the DRF-3SLS forecasts. the GRF forecasts slightly outperform 3SLS and FIML in terms of forecast error variance. The mean of GRF forecasts lies between the means of ULS forecasts and DRF-3SLS forecasts, and it becomes closer to DRF-3SLS forecasts as a smaller value of c is assessed for less uncertain structural specification. Apparantly the DRF-2SLS forecasts perform worst

[Table 3] Experiment I :
Medium sample ($T=48$), higher centrality, correct specifidation

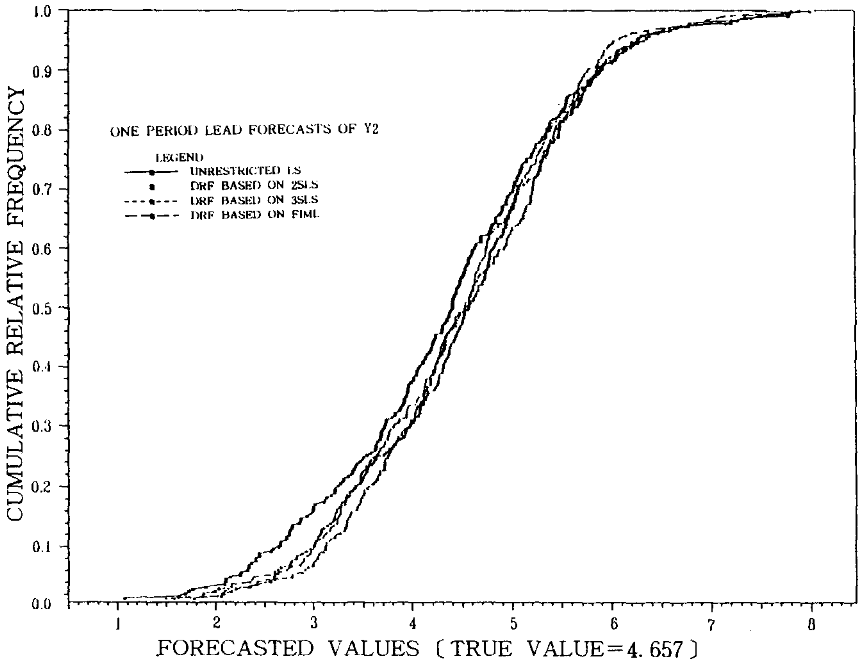
	mean	st. dv.	skewness	kurtosis	median	length of interval 50%	80%	100%
Model A, one-period lead forecasts of $Y_1=2.578$, one-degree overidentified								
ULS	2. 501	. 346	-. 01	-. 13	2. 51	. 46	. 93	1. 92
DRF-2SLS	2. 548	. 304	. 03	. 03	2. 55	. 40	. 79	1. 62
DRF-3SLS	2. 568	. 296	. 03	. 18	2. 57	. 40	. 75	1. 73
DRF-FIML	2. 537	. 295	-. 01	. 09	2. 55	. 41	. 75	1. 69
PRRF	2. 467	. 375	-. 18	. 21	2. 48	. 47	. 97	2. 20
MSRF	2. 559	. 295	. 03	. 11	2. 57	. 40	. 76	1. 70
GRF ($c=1.0$)	2. 531	. 292	. 01	. 09	2. 55	. 41	. 73	1. 68
GRF ($c=.05$)	2. 534	. 291	. 02	. 09	2. 54	. 41	. 74	1. 68
Model A, one-period lead forecasts of $Y_2=4.657$, one-degree overidentified								
ULS	4. 408	1. 214	. 07	. 07	4. 46	1. 53	3. 15	6. 76
DRF-2SLS	4. 571	1. 119	. 12	. 05	4. 62	1. 54	2. 77	5. 98
DRF-3SLS	4. 635	1. 077	. 13	. 20	4. 59	1. 50	2. 70	6. 37
DRF-FIML	4. 520	1. 077	. 06	. 09	4. 58	1. 55	2. 67	6. 26
PRRF	4. 384	1. 296	. 06	-. 11	4. 44	1. 72	3. 31	7. 27
MSRF	4. 602	1. 078	. 11	. 15	4. 58	1. 49	2. 69	6. 32
GRF ($c=1.0$)	4. 501	1. 067	. 09	. 13	4. 54	1. 39	2. 67	6. 22
GRF ($c=.05$)	4. 509	1. 064	. 10	. 11	4. 54	1. 47	2. 66	6. 21
Model B, one-period lead forecasts of $Y_1=5.047$, just identified								
ULS	5. 369	. 344	. 17	. 20	5. 36	. 43	. 89	2. 17
DRF-2SLS	5. 387	. 310	. 18	. 05	5. 38	. 40	. 80	1. 76
DRF-3SLS	5. 405	. 301	. 10	. 03	5. 39	. 43	. 75	1. 69
DRF-FIML	5. 393	. 302	. 09	. 02	5. 39	. 42	. 76	1. 67
PRRF	5. 369	. 344	. 17	. 20	5. 36	. 43	. 89	2. 17
MSRF	5. 405	. 301	. 10	. 03	5. 39	. 43	. 75	1. 69
GRF ($c=1.0$)	5. 390	. 300	. 14	. 06	5. 38	. 41	. 76	1. 71
GRF ($c=.05$)	5. 394	. 302	. 09	. 02	5. 59	. 42	. 76	1. 67
Model B, one-period lead forecasts of $Y_2=21.629$, two-degree overidentified								
ULS	21. 490	1. 210	. 22	. 36	21. 40	1. 56	3. 07	7. 61
DRF-2SLS	18. 383	1. 155	. 19	. 22	21. 49	1. 46	2. 97	6. 78
DRF-3SLS	21. 597	1. 120	. 10	. 21	21. 58	1. 46	2. 80	6. 50
DRF-FIML	21. 563	1. 124	. 08	. 19	21. 58	1. 48	2. 74	6. 43
PRRF	21. 452	1. 299	. 20	. 36	21. 38	1. 68	3. 31	8. 52
MSRF	21. 597	1. 120	. 10	. 21	21. 58	1. 46	2. 80	6. 50
GRF ($c=1.0$)	21. 554	1. 115	. 13	. 24	21. 51	1. 38	2. 83	6. 54
GRF ($c=.05$)	21. 564	1. 123	. 09	. 19	21. 58	1. 45	2. 76	6. 44



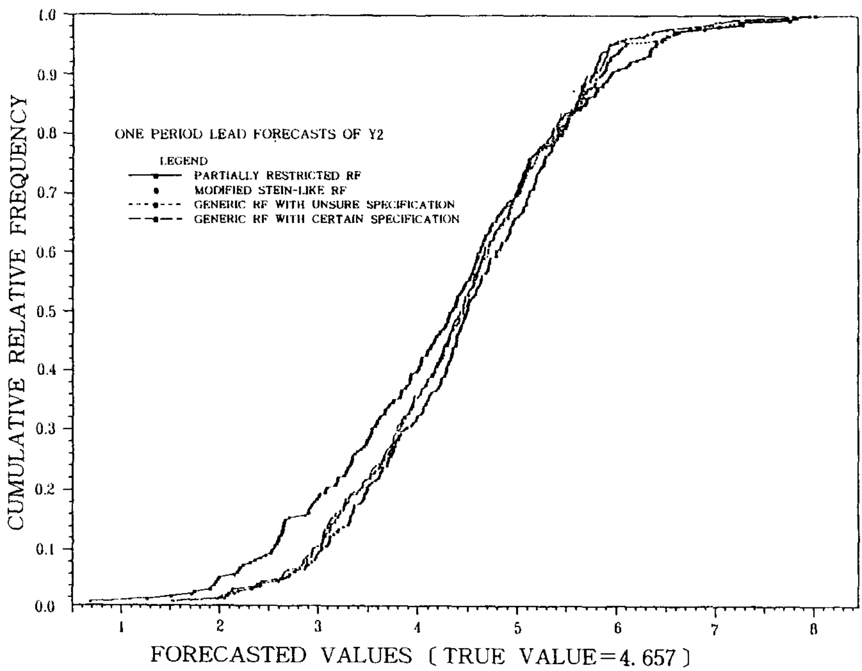
[Figure 1] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)



[Figure 2] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)



[Figure 3] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)



[Figure 4] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)

among efficient SEM estimators, and the ULS forecasts perform worst among all standard SEM estimators, because they ignore some or all of the structural information under correct specification. Under correct specification, ULS is not a recommendable forecasting method, since it ignores structural information and that causes a substantial loss of efficiency. Surprisingly, the PRRF forecasts do not outperform either the ULS forecasts or the DRF-2SLS forecasts, even though PRRF incorporates more of the correct structural information than the ULS and it possesses all finite moments. This happens in the situation where the experimental design satisfies Nagar and Sahay's (1978, p 234) requirement that the absolute value of the included other endogenous variable is less than one (.5 in the first equation), Figures 1 through 4 demonstrate that the ULS forecasts behave differently from the standard SEM forecasts, and that the PRRF forecasts deviate from the MSRF or GRF forecasts which perform as well as the DRF-FIML or DRF-3SLS forecast, for all Y_1 and Y_2 in the models of different overidentifying degrees.

2. Distributions of Forecasts Under Small Samples

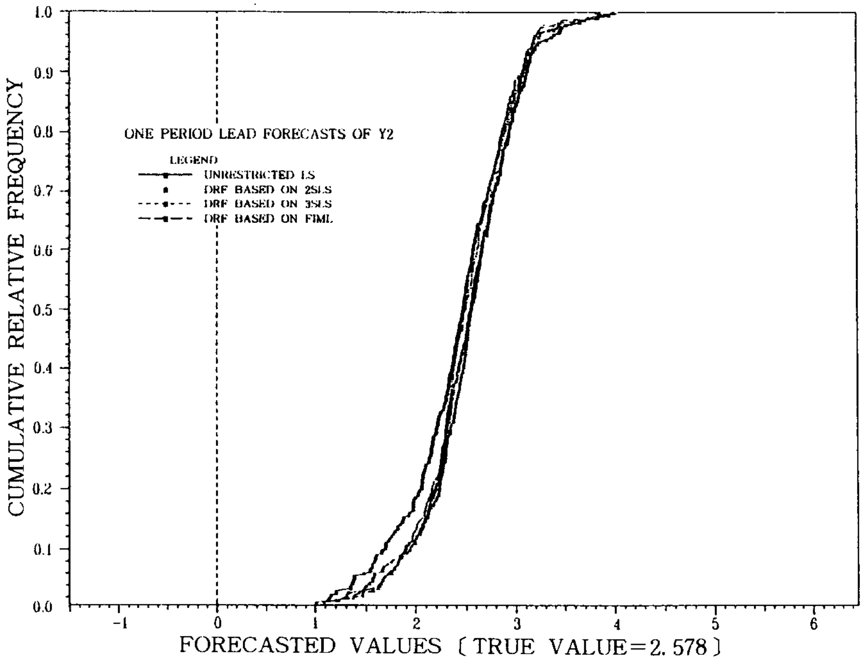
Table 4 and Figures 5 through 8 report the results from Experiment Ⅱ. Compared to Experiment Ⅰ, the risk (in terms of standard deviation of the samples) is increased overall by around 60%, when the sample size is halved. With reduced sample size, the length of the 80% central interval (inter-quintile range) is increased by about 45%, while that of the whole range of the samples is increased by about 75% in general. This indicates that the tail areas of almost all the forecasts become much fatter. Therefore we are more likely to produce forecasts which deviate significantly from the true values. When the sample size is halved, the differences between the DRF-FIML and DRF-3SLS forecasts and the MSRF and GRF forecasts are still negligible. However, the PRRF forecasts exhibit significantly fatter lower tail areas. See Figures 10 and 12. Even so, the general observations of Experiment I about the relative performance of the different forecasts are maintained in small sample situations. 'Steinian' issues related to the statistical improvement of a positive-part MSRF prediction upon traditional DRF-3SLS forecasts in small samples, are not observed in this experiment. This may be due to the design of the MSRF estimator which becomes almost identical to DRF-3SLS under correct specification and at usual significance level.

3. Distributions of Forecasts Under Departures from Normality

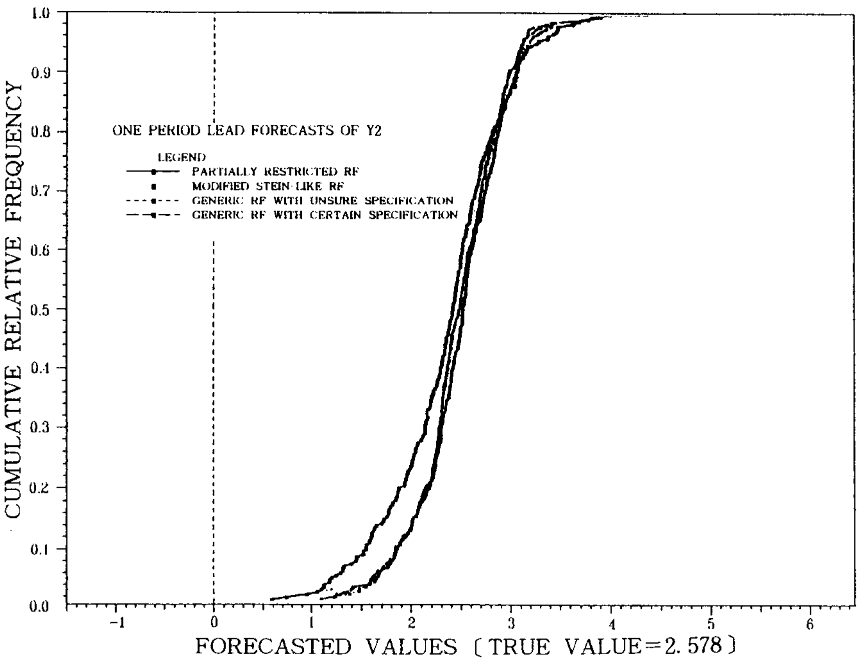
Table 5 and Figures 9 through 16 report the results from Experiment Ⅲ. The sample second moment matrix of exogenous variables, $Z'Z$, is changed from $24\,I_4$ in Experiment Ⅱ to I_4 in Experiment Ⅲ. The value of the concentration parameter for the system, $P'Z'ZP/T$, and hence, the sample variability matrix $\Omega^{-1}P'Z'ZP/T$

[Table 4] Experiment Ⅱ :
Small sample ($T=24$), higher centrality, correct specification

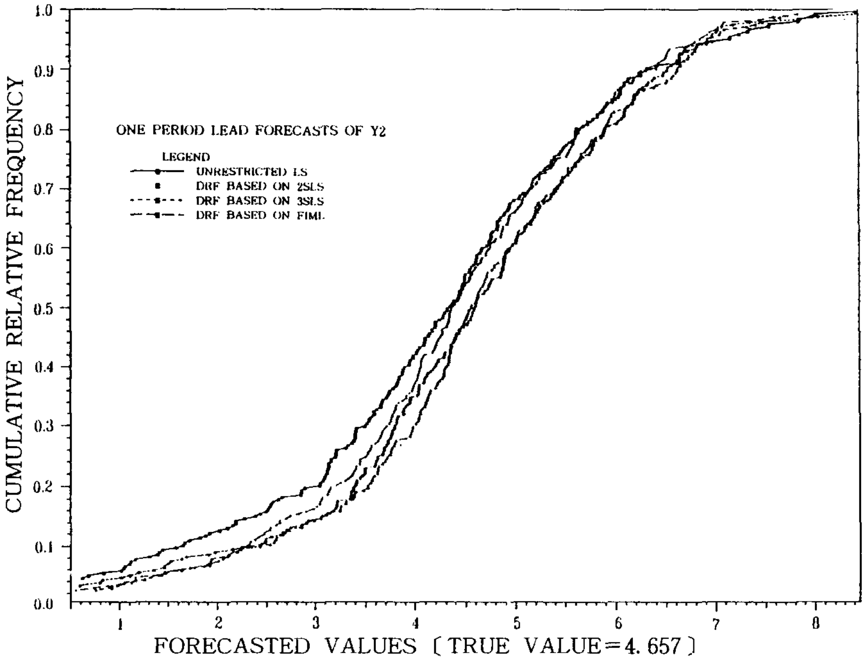
	mean	st. dv.	skewness	kurtosis	median	length of interval 50%	80%	100%
Model A, one-period lead forecasts of $Y_1=2.578$, one-degree overidentified								
ULS	2.479	.519	-.18	.28	2.50	.63	1.33	2.93
DRF-2SLS	2.562	.457	-.27	.81	2.58	.54	1.12	2.81
DRF-3SLS	2.586	.443	-.15	.84	2.58	.53	1.11	2.83
DRF-FIML	2.529	.436	-.05	1.04	2.53	.52	1.08	2.95
PRRF	2.412	.573	-.26	.35	2.46	.67	1.47	3.38
MSRF	2.553	.446	-.11	.70	2.56	.55	1.16	2.84
GRF ($c=1.0$)	2,517	.440	-.14	.91	2.53	.50	1.07	2.80
GRF ($c=.05$)	2.522	.432	-.16	.89	2.53	.52	1.08	2.75
Model A, one-period lead forecasts of $Y_2=4.657$, one-degree overidentified								
ULS	4.332	1.834	-.22	.45	4.43	2.19	4.76	11.14
DRF-2SLS	4.636	1.687	-.34	.74	4.65	1.99	4.05	10.13
DRF-3SLS	4.704	1.631	-.21	.77	4.69	1.93	4.06	10.21
DRF-FIML	4.491	1.612	-.15	.84	4.47	1.99	3.98	9.89
PRRF	4.305	1.936	-.22	.22	4.39	2.32	5.07	11.16
MSRF	4.584	1.640	-.19	.62	4.64	1.94	4.29	10.27
GRF ($c=1.0$)	4.451	1.619	-.23	.81	4.46	1.89	4.03	9.97
GRF ($c=.05$)	4.463	1.593	-.25	.78	4.43	1.90	4.00	9.86
Model B, one-period lead forecasts of $Y_1=5.047$, just identified								
ULS	5.308	.519	-.18	.28	5.33	.63	2.50	2.93
DRF-2SLS	5.339	.461	-.09	.71	5.35	.55	1.33	2.86
DRF-3SLS	5.372	.447	.04	.73	5.37	.49	1.14	2.76
DRF-FIML	5.352	.446	.04	.76	5.36	.50	1.14	2.77
PRRF	5.308	.519	-.18	.28	5.33	.63	1.11	2.93
MSRF	5.372	.447	.04	.73	5.37	.49	1.33	2.76
GRF ($c=1.0$)	5.345	.447	.01	.70	5.36	.51	1.14	2.80
GRF ($c=.05$)	5.353	.447	.05	.76	5.36	.51	.1.14	2.77
Model B, one-period lead forecasts of $Y_2=21.629$, two-degree overidentified								
ULS	21.304	1.834	-.22	.45	21.41	2.19	4.76	11.14
DRF-2SLS	21.358	1.733	-.20	.71	21.45	1.99	4.36	10.99
DRF-3SLS	21.490	1.674	-.08	.74	21.54	1.72	4.42	10.59
DRF-FIML	21.434	1.674	-.07	.77	21.44	1.73	4.31	10.60
PRRF	21.231	1.978	-.26	.26	21.34	2.54	5.13	11.25
MSRF	21.490	1.674	-.08	.74	21.54	1.72	4.42	10.59
GRF ($c=1.0$)	21.413	1.671	-.10	.74	21.48	1.75	4.31	10.69
GRF ($c=.05$)	21.439	1.675	-.07	.78	21.46	1.71	4.28	10.60



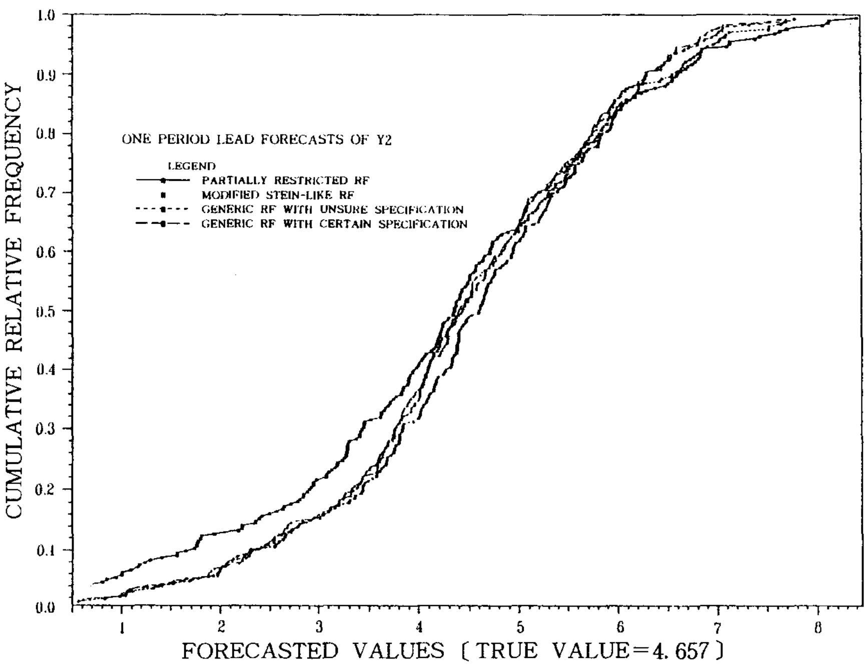
[Figure 5] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)



[Figure 6] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 24)



[Figure 7] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)



[Figure 8] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, CORRECT SPECIFICATION, SAMPLE SIZE 48)

becomes smaller in magnitude (in terms of its determinant). It is well known in exact density theory that the smaller value of the non-centrality parameter (means-sigma matrix) adversely affects the p. d. f. of the instrumental variable estimators. Even though the sample size remains unchanged compared to Experiment Ⅱ, the lower-centrality affects the DRF-2SLS and DRF-3SLS forecasts substantially (Basmann (1963)). Especially in model A in Experiment Ⅲ, where all equations are overidentified by one degree, the interval of the whole samples becomes almost infinitely wider than the inter-quintile interval for the standard reduced form forecasts ! (about 280 times wider for DRF-2SLS, 200 times for DRF-3SLS, and 40 times for DRF-FIML forecast) A serious outlier problem still remains in model B for DRF-2SLS and DRF-3SLS but to a much lower degree. The MSRF forecasts are specific to the specification test methods employed, because, for the same disturbance structures, the value of λ in the MSRF forecast in model A is almost equal to 1 (i. e., MSRF becomes identical to DRF-3SLS), while it is a great deal less than 1 in model B such that outliers are removed. The size of the whole sample range interval is 230 times wider than that of the inter-quintile interval for the MSRF forecast in model A, while it is only 2. 5 times wider in model B. This indicates that the local power of the specification test is specific to the model parameters. In this sense it is important to compare alternative forecasts from different estimation techniques in the same spirit as Hausman's (1978) specification test.

The empirical c. d. f. 's of the DRF-2SLS forecast and the DRF-3SLS forecasts in Figures 9 and 13 (and, for some cases, those of the MSRF forecasts in Figures 10 and 14) suggest that these forecasts cause unbounded risk in finite samples and under lower centrality. Contrary to traditional emphasis on the sample size in asymptotic theory, the concentration parameter in terms of the sample second moment matrix of exogenous variables is a more important factor in detecting departures from normality. The lower centrality seriously affects the distribution of estimators and forecasts as evidenced by the significant outliers in Experiment Ⅲ compared to Experiment Ⅱ. The third and fourth sample moments are far from zero. In this situation, the GRF forecasts and the DRF-FIML forecasts outperform all other forecasts in most cases. The ULS and PRRF forecasts become relatively better under the erratic behavior of DRF-2SLS and DRF-3SLS. 'Outlier frequencies' at both ends of the empirical p. d. f. in Figures 11 and 12 for Y_1 and Figures 15 and 16 for more dispersely distributed Y_2 clearly demonstrate the seriousness of the departure from normality in these experiments.

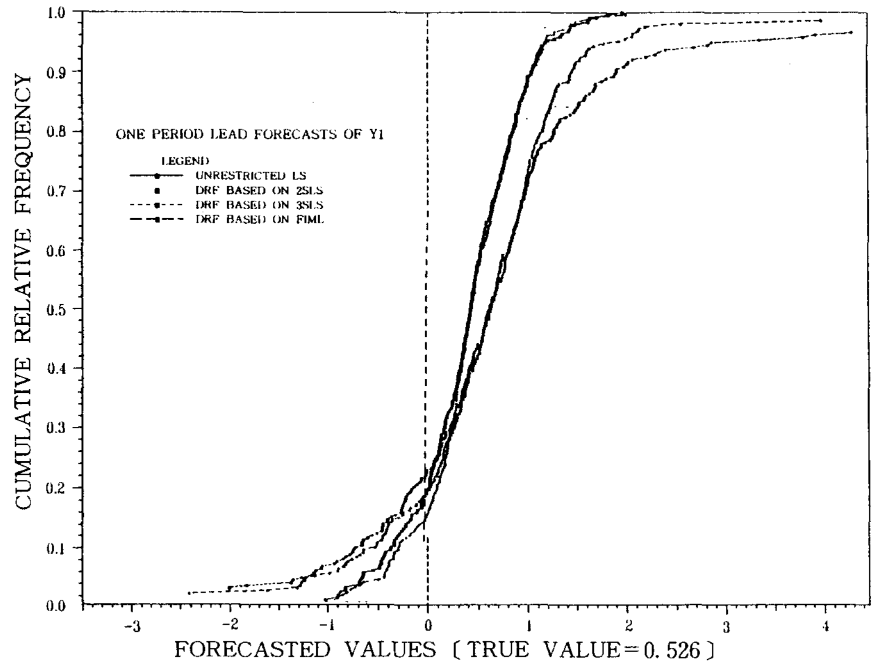
4. Distributions of Forecasts Under Exclusion-Inclusion Misspecifications

Table 6 and Figures 17 through 20 report the results from Experiment Ⅳ. The structural misspecification considered here is of the exclusion and inclusion type. Under serious misspecification, the ULS forecasts perform consistently well throughout

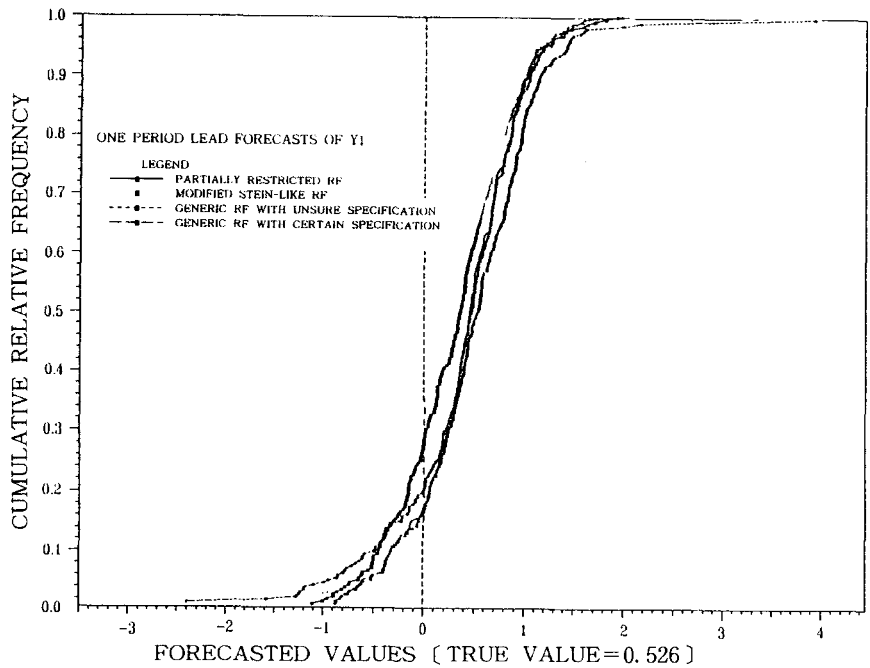
[Table 5] Experiment Ⅲ :

Small sample (T=24), higher centrality, correct specifidation

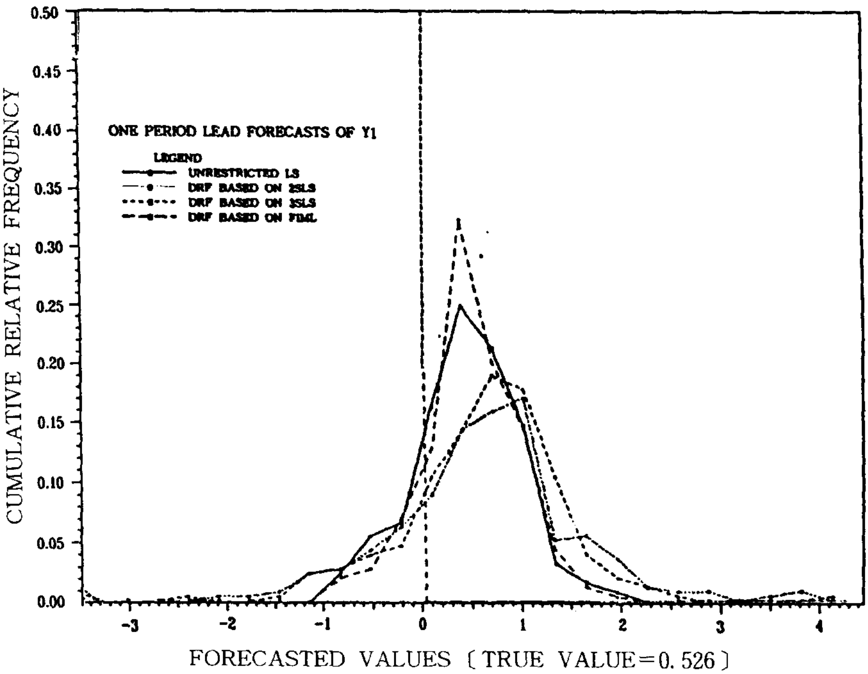
						length of interval		
	mean	st. dv.	skewness	kurtosis	median	50%	80%	100%
Model A, one-period lead forecasts of Y_1=.526, one-degree overidentified								
ULS	.427	.519	-.18	.28	.45	.63	1.33	2.93
DRF-2SLS	4.082	44.523	14.67	222.83	.64	.92	2.52	706.02
DRF-3SLS	2.223	24.159	15.62	245.84	.65	.84	1.91	387.52
DRF-FIML	.648	3.099	15.26	238.33	.44	.55	1.28	49.77
PRRF	.321	.557	.017	.003	.36	.70	1.46	3.04
MSRF	2.037	24.125	15.707	247.728	.54	.75	1.62	387.52
GRF (c=1.0)	.439	.469	-.319	.371	.47	.60	1.30	2.86
GRF (c=.05)	.438	.505	-.250	.272	.45	.62	1.35	2.87
Model A, one-period lead forecasts of Y_2=0.951, one-degree overidentified								
ULS	.626	1.834	-.22	.45	.73	2.19	4.76	11.14
DRF-2SLS	12.818	147.477	14.83	227.03	1.42	3.09	8.49	2345.45
DRF-3SLS	6.748	80.421	15.62	245.90	1.46	2.83	6.39	1289.22
DRF-FIML	1.372	10.250	15.67	236.54	.75	2.03	4.55	164.89
PRRF	.63471.853		-.23	.39	.73	2.28	4.87	11.08
MSRF	6.024	80.294	15.72	248.06	1.00	2.55	5.69	1288.34
GRF (c=1.0)	.684	1.767	-.40	.51	.76	2.09	4.73	10.73
GRF (c=.05)	.678	1.796	-.33	.37	.78	2.24	4.82	10.72
Model B, one-period lead forecasts of Y_1=1.104, just identified								
ULS	1.004	.519	-.18	.28	1.03	.63	1.33	2.93
DRF-2SLS	.898	2.534	-14.05	212.53	1.09	.60	1.35	40.95
DRF-3SLS	1.094	.744	-7.09	82.73	1.13	.47	1.17	10.44
DRF-FIML	1.036	.440	.05	.73	1.05	.50	1.11	2.79
PRRF	1.004	.519	-.18	.28	1.03	.63	1.33	2.93
MSRF	1.132	.461	-.11	.97	1.13	.47	1.17	3.03
GRF (c=1.0)	1.044	.443	.00	.77	1.04	.47	1.11	2.88
GRF (c=.05)	1.045	.446	-.06	.80	1.06	.48	1.13	2.89
Model B, one-period lead forecasts of Y_2=4.415, two-degree overidentified								
ULS	4.090	1.834	-.22	.45	4.19	2.19	4.76	11.14
DRF-2SLS	3.513	10.116	-14.20	215.83	4.28	2.16	5.29	164.07
DRF-3SLS	4.289	2.908	-7.60	90.74	4.46	1.81	4.65	41.31
DRF-FIML	4.175	1.654	-.12	.77	4.22	1.83	4.18	10.68
PRRF	3.933	1.985	-.29	.25	4.03	2.50	5.11	11.20
MSRF	4.444	1.733	-.24	.93	4.45	1.78	4.63	11.00
GRF (c=1.0)	4.198	1.658	-.15	.81	4.23	1.79	4.24	10.77
GRF (c=.05)	4.191	1.674	-.20	.81	4.22	1.77	4.26	10.70



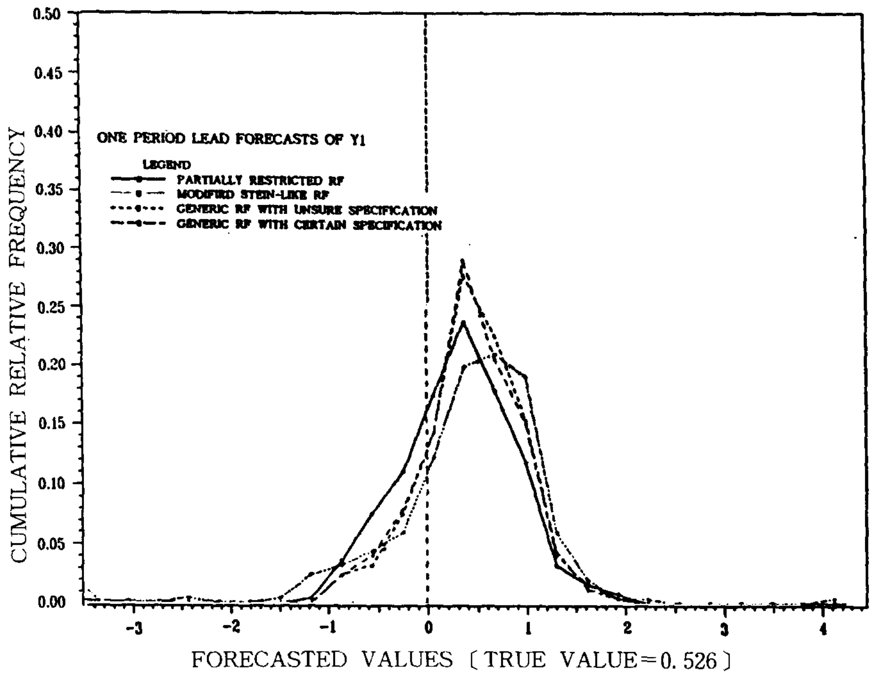
[Figure 9] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



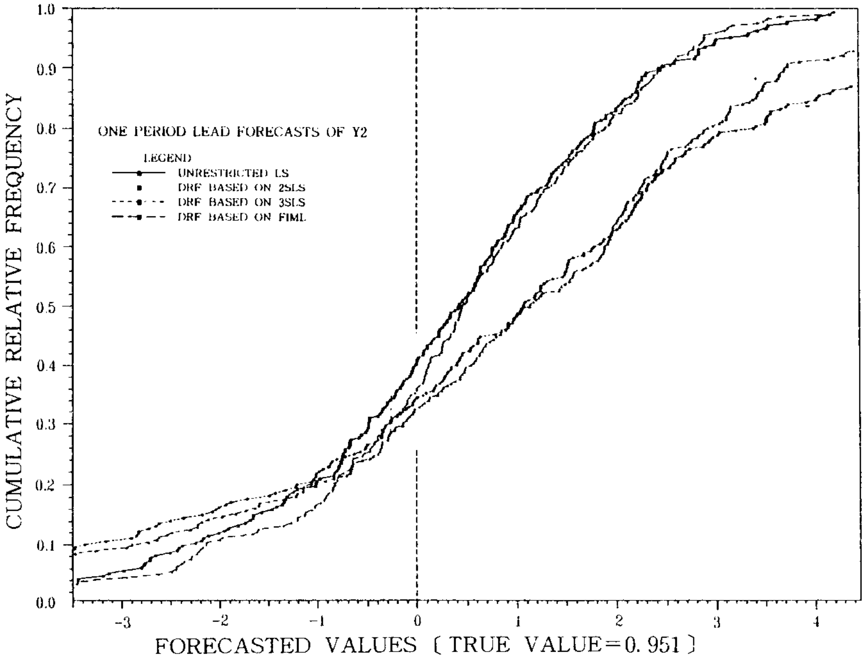
[Figure 10] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



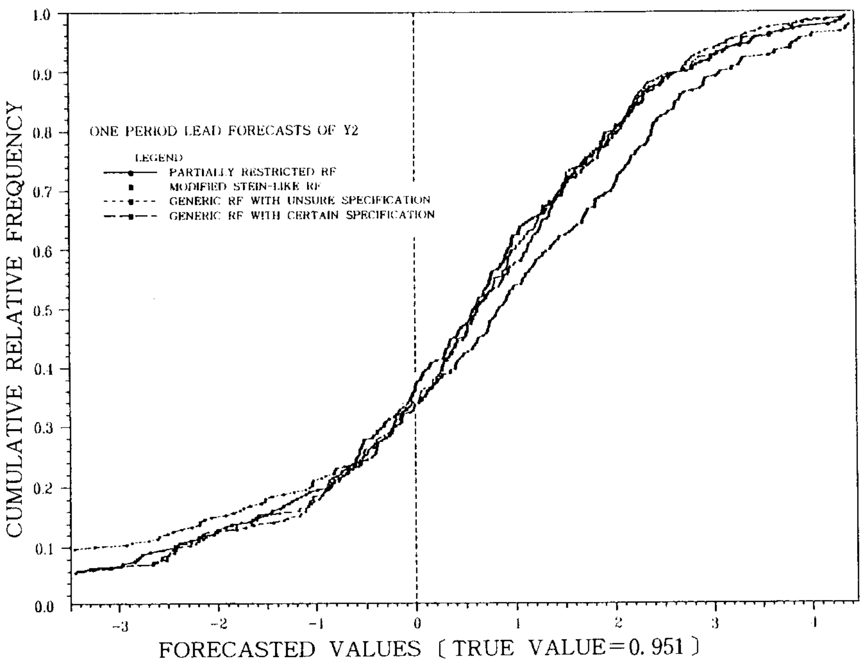
[Figure 11] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS
(BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



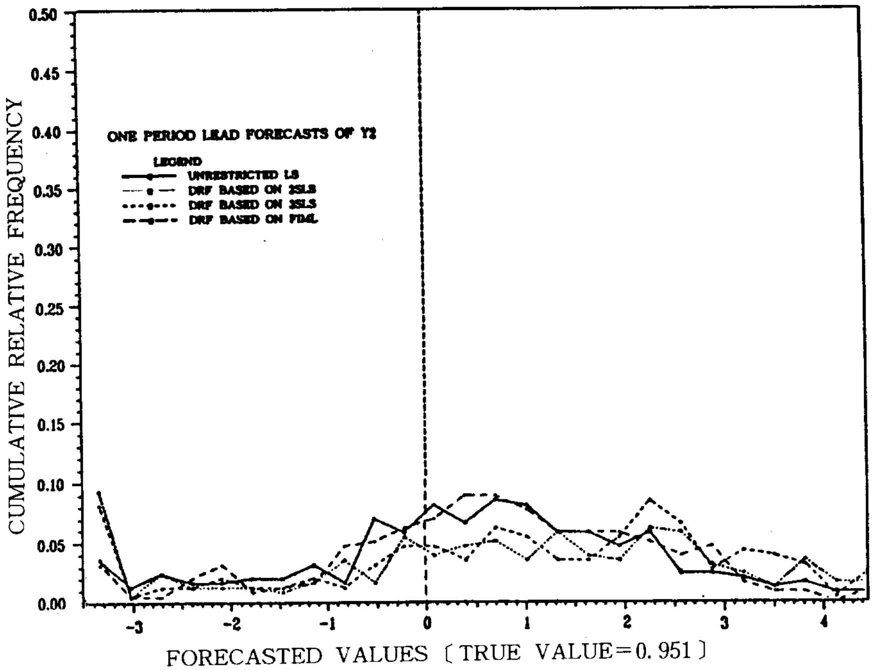
[Figure 12] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS
(BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



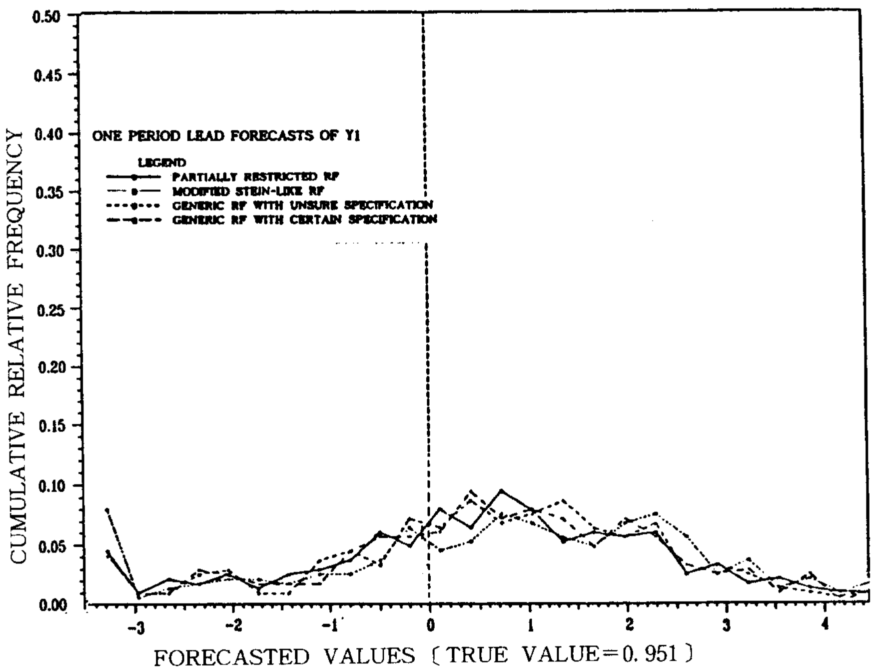
[Figure 13] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



[Figure 14] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



[Figure 15] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)



[Figure 16] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL A, LOWER CENTRALITY, SAMPLE SIZE 24)

the model because ULS does not incorporate any structural identifying restrictions. As a mixture of ULS and structural 2SLS, the PRRF forecasts also have nice distributions which are very similar to those of the ULS forecast. The distributions of the DRF-3SLS and DRF-2SLS forecasts are the worst among predictors because of the non-existence of moments of underlying derived reduced form estimators. In small samples these unbounded moments are seriously aggravated by the structural misspecification as demonstrated in our experimental models which has four endogenous variables with varieties of identification degrees. Unbounded risk is obvious in these situations. Even though the DRF-FIML forecasts have finite moments, DRF-FIML incorporates structural specification constraints in an exact manner, hence, the resulting forecasts are sensitive to structural misspecification. See Figures 17 and 19.

The MSRF forecasts exhibit contradictory behavior in model A and model B. As expected from the previous discussions, when the Wald type specification test successfully rejects the wrongful restrictions, they outperform all alternative forecasts (Figures 18 and 20). But, when it cannot reject them for a given significance level, they behave just as poorly as the DRF-3SLS forecasts (the case of model A in Table 6). In terms of the R. M. S. E. criterion, the MSRF forecasts have the lowest risk in model B, but as high a risk as the DRF-3SLS forecasts in model A.

Under serious misspecification, the GRF forecasts do not provide distributions which are as 'good' as those of ULS or PRRF forecasts. However, the GRF forecasts do provide better predictive distributions than any other of the efficient structural SEM estimators studied. In fact, a model builder's assessed value for c as a measure of prior structural specification uncertainty does not change the performance of the GRF forecasts very much as is clear from the comparison of Figures 18, and 20. This confirms our earlier observations from the distributions of the GRF reduced form coefficient estimates in Maasoumi and Jeong (1988) and Jeong (1985).

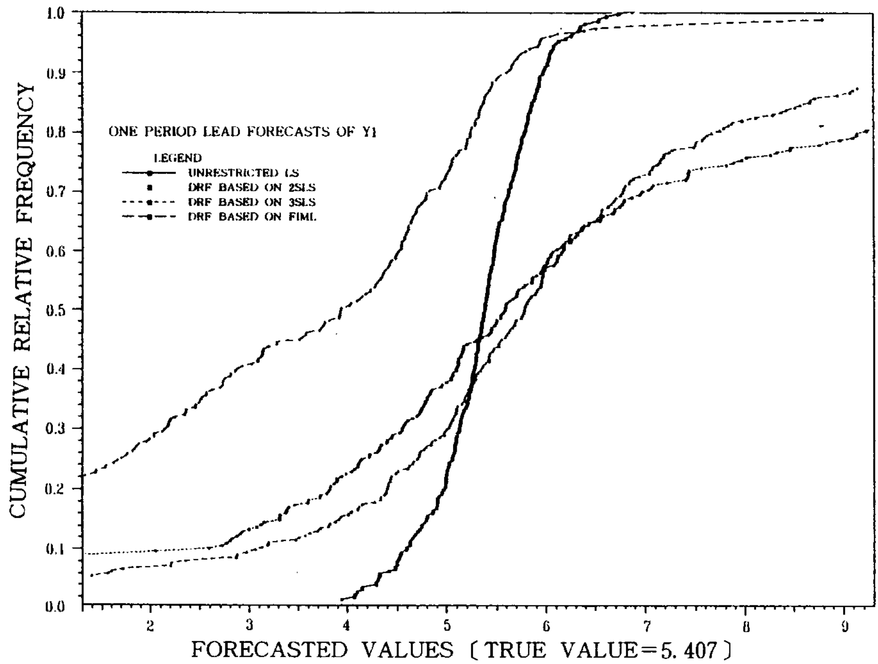
V. SUMMARY AND CONCLUSION

As was emphasized at the beginning of Section 2, our goal in this Monte Carlo study is to find a 'good' predictive distribution, one that perform reasonably well relative to other alternative forecasts in various experimental situations controlled in terms of the model parameter changes. In Experiments I and II, standard efficient SEM estimators provide reasonable forecasts. However, in Experiments III and IV, where the reduced form coefficient estimators are seriously affected by the sample second moment matrices and the correctness of the imposed specification constraints, the derived reduced form estimators without finite moments tend to make bad

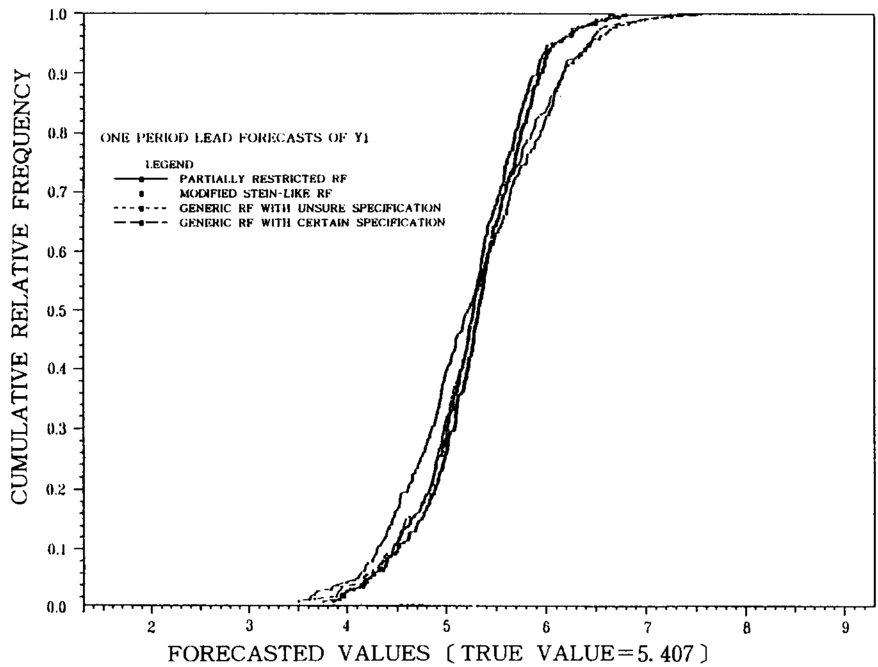
[Table 6] Experiment Ⅳ :
Small sample ($T=24$), higher centrality, misspecification

	mean	st. dv.	skewness	kurtosis	median	length of interval 50%	80%	100%
Model A, forecasts of $Y_1=2.578$, $v_1=1$ misspecified as $v_1=0$								
ULS	2.479	.519	-.18	.28	2.50	.63	1.33	2.93
DRF-2SLS	-1.027	44.819	.14	56.15	.44	5.09	25.67	779.69
DRF-3SLS	1.010	25.321	2.31	59.44	1.67	3.19	13.99	461.85
DRF-FIML	2.428	1.390	-.55	.67	2.73	1.50	3.95	8.45
PRRF	2.479	.519	-.18	.28	2.50	.63	1.33	2.93
MSRF	1.040	25.322	1.31	59.43	1.93	3.16	13.99	461.85
GRF ($c=1.0$)	2.983	.678	.03	-.47	3.01	1.00	1.76	3.43
GRF ($c=.05$)	2.981	.743	.12	4.90	2.93	1.11	1.95	3.81
Model A, forecasts of $Y_2=4.657$, $v_2=1$ misspecified as $v_2=2$								
ULS	4.332	1.834	-.22	.45	4.43	2.19	4.76	11.14
DRF-2SLS	-2.257	106.369	.36	57.58	1.08	12.40	57.33	1882.57
DRF-3SLS	2.406	60.853	1.35	60.08	4.17	7.65	33.02	1115.13
DRF-FIML	4.750	4.452	-.60	.29	5.41	5.25	12.56	24.30
PRRF	5.855	1.390	.17	.31	5.80	1.62	3.57	8.00
MSRF	2.212	60.845	1.36	60.14	3.81	7.25	33.02	1115.13
GRF ($c=1.0$)	6.186	2.455	-.06	-.48	6.37	3.52	6.31	12.39
GRF ($c=.05$)	6.186	2.455	-.06	-.48	6.37	3.52	6.31	12.39
Model B, forecasts of $Y_2+5.407$, $v_2=0$ misspecified as $v_1=1$								
ULS	5.308	.519	-.18	.28	5.33	.63	1.33	2.93
DRF-2SLS	15.109	117.654	15.18	236.15	5.34	4.07	10.80	1906.14
DRF-3SLS	12.484	88.047	15.26	238.17	5.74	2.47	7.12	1450.06
DRF-FIML	3.501	6.103	.40	58.28	3.88	3.35	5.68	109.49
PRRF	5.370	.519	-.20	.30	5.39	.66	1.34	2.87
MSRF	5.314	.517	-.17	.26	5.34	.62	1.37	2.94
GRF ($c=1.0$)	5.328	.707	.16	-.21	5.28	1.01	1.76	3.76
GRF ($c=.05$)	5.384	.637	.31	.85	5.34	.79	1.60	4.34
Model B, forecasts of $Y_2=21.629$, $v_2=2$ misspecified as $v_2=1$								
ULS	21.304	1.834	-.22	.45	21.41	2.19	4.76	11.14
DRF-2SLS	59.248	456.966	15.16	235.74	22.44	15.05	43.41	7404.70
DRF-3SLS	49.033	341.950	15.25	237.80	23.00	9.44	28.73	5634.29
DRF-FIML	14.662	23.727	.40	58.59	16.23	12.92	21.39	424.91
PRRF	21.277	1.936	-.22	.22	21.36	2.32	5.09	11.16
MSRF	21.325	1.830	-.22	.43	21.45	2.16	4.79	11.19
GRF ($c=1.0$)	21.363	2,332	.11	.01	21.27	3.09	6.09	12.90
GRF ($c=.05$)	21.558	2.154	.25	1.00	21.46	2.73	5.62	14.79

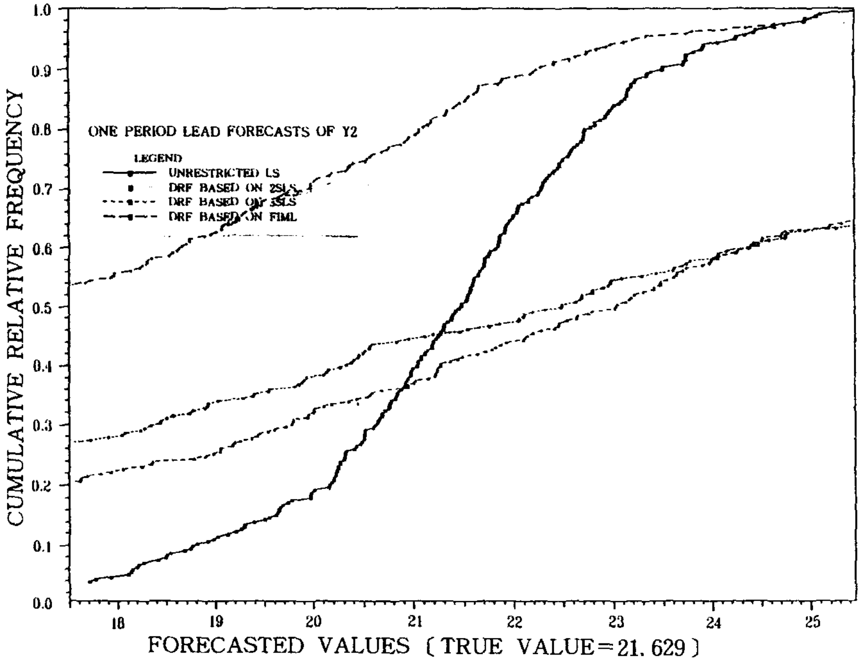
** V_1 (V_2) = the degree of overidentification of the first (second) equation



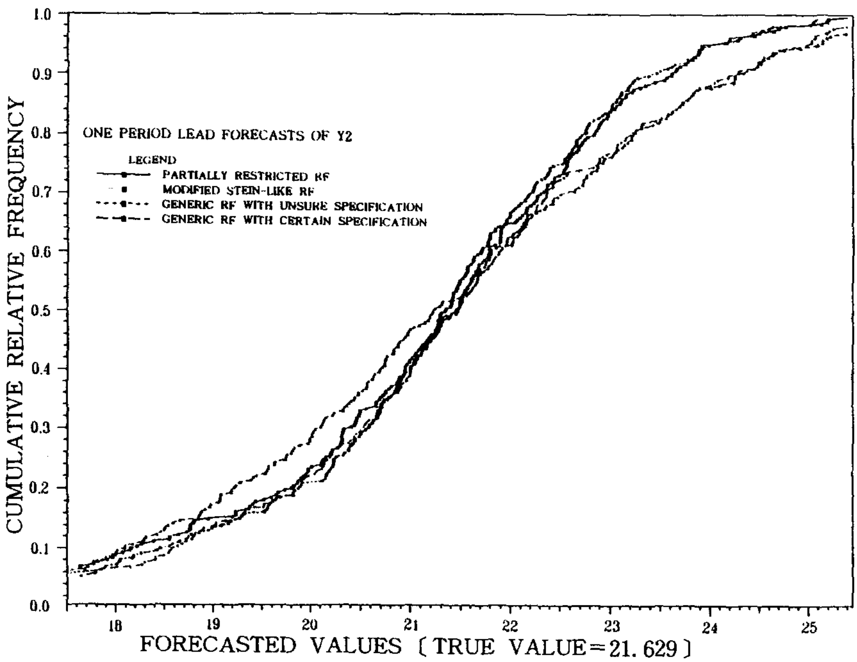
[Figure 17] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL B, MISSPECIFIED AS MODEL A, SAMPLE SIZE 24)



[Figure 18] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL B, MISSPECIFIED AS MODEL A, SAMPLE SIZE 24)



[Figure 19] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL B, MISSPECIFIED AS MODEL A, SAMPLE SIZE 24)



[Figure 20] EMPIRICAL CUMULATIVE DISTRIBUTIONS FROM 250 REPLICATIONS (BASE MODEL B, MISSPECIFIED AS MODEL A, SAMPLE SIZE 24) forecasts exhibiting unbounded risks, as shown by fatter tail areas. The contrasts between the traditional SEM forecasts (Figure 17 or 19, for example) and the lower risk reduced form forecasts (Figure 18 or 20 respectively) demonstrate how important the choice of predictive distributions is under general quadratic risk criteria. With identical data set and under the same structural specifications, the choice of the reduced form estimator matters for risk involved in forecasting. The conclusion about the choice of predictive distribution may be summarized as follows based on our Monte Carlo simulations.

1) The GRF forecast with 'unsure' structural specification outperforms all the existing SEM estimators (except for the DRF-FIML forecast for some cases) under general quadratic risk measures. As a matrix weighted average of ULS and DRF-3SLS forecasts, it mixes more prior information under correct specification, and it mixes more sample information under structural misspecification and with erratic behaviors (infinite moments) in DRF-3SLS. Hence it provides reasonable predictive distributions in almost all situations.

2) The DRF-FIML forecast is the only standard efficient SEM forecast which has some finite moments. Contrary to traditional beliefs in analytic studies, it performs very well in small samples under correct specification, and reasonably well under low centrality and even under misspecification. The reasonable performance of FIML predictive distribution may be due to the existence of some finite moments and the higher centrality which favors normality assumption.

3) The MSRF forecast does not perform as well as expected as a member of the class of 'Steinian' estimators. Especially in Model A where both equations are overidentified by one degree, the Wald type asymptotic specification test employed found to be not powerful enough to discriminate the erratic behavior of DRF-3SLS reduced form. In this situation, the MSRF forecast behaves just like DRF-3SLS forecast. However, in model B where one equation is just identified and the other equation is overidentified by two degrees, the MSRF predictive distribution optimally combines ULS and DRF-3SLS predictive distributions and outperforms all the traditional SEM forecasts. In this setting, the MSRF forecast even outperform the GRF forecast in some cases.

4) The DRF-3SLS forecast is asymptotically equivalent to the DRF-FIML forecast under correct specification and higher centrality. In this situation, it performs as well as DRF-FIML ever with non-existence of moments. This Monte Carlo study indicates that the problem that forces erratic behavior of the DRF-3SLS forecast is the non-existence of moments and that this abnormality is mostly affected by the departure from normality due to either the lower sample second moment matrix of exogenous variables or the incorrect coefficient constraints.

5) The ULS forecast ignores the structural coefficient constraints defined in (9). When these constraints are true it suffers loss of efficiency, but, when these constraints are incorrect, it is benifited by ignoring such information and it out performs those estimators that are seriously affected by this misspecification.

6) The PRRF forecast mixes the ULS estimates and the structural-2SLS estimates not in an optimal way but in an approximate way. Its performence is worst among all SEM estimators under correct specification and/or with higher centrality. However, it does perform relatively well under misspecification for the same reason that the ULS forecast does.

7) The DRF-2SLS forecast exhibits the worst performance in this experimental study. This is an important result when we consider its popularity amongst applied macro economists despite a decade of warnings from small sample theorists.

Our experiments, as is the case with all Monte Carlo studies, are limited to representing only a small set of parameter space in which the SEM resides. Diversification of the identifiability conditions restricts the number of endogenous and exogenous variables involved in our experimental design. This may be unfavorable for some estimators and predictors, especially for MSRF, since the preliminary test employed has little power to discern just-identified misspecification. However we will leave more detailed exploration of the properties of these predictors in more diverse situations for future study.

REFERENCES

[1] AITCHISON, J. (1975) "Goodness of prediction fit," *Biometrika*, 62, 547—554.

[2] AKAIKE, H. (1977), "An objective use of Bayesian models," *Annals of the Institute of Statistics and Mathematics*, 29, 9—20.

[3] ANDERSON, T. W., K. MORIMUNE and T. SAWA (1983), "The numerical values of some key parameters in econometric models," *Journal of Econometrics*, 21, 229—243.

[4] BASMANN, R. L. (1963), "A note on the exact finite sample frequency functions of generalized classical linear estimators in a leading three equation case," *Journal of the American Statistical Association*, 58, 161—171.

[5] CHAMBERLAIN, G. and E. LEAMER (1976), "Matrix Weighted averages and posterior bounds," *Journal of the Royal Statistical Society*, B38, 73—84.

[6] DHRYMES, P. (1973), "Restricted and unrestricted reduced forms : asymptotic distribution and relative efficiency," *Econometrica*, 41, 119—134.

[7] GOLDBERGER, A. S., A. L. NAGAR and H. S. ODEH (1961), "The covariance matrices of reduced form coefficients and forecasts for a structural econometric model," *Econometrica*, 29, 556-573.

[8] HAUSMAN, J. A. (1975), "An instrumental variable approach to full information estimators for linear and certain nonlinear econometric method," *Econometrica*, 43, 727—738.

[9] HAUSMAN, J. A. (1978), "Specification tests in econometrics," *Econometrica*, 46, 1251—1272.

[10] JEONG, JIN-HO (1985), "Lower Risk Reduced Forms Estimation and Forecasting Based on Finite Observations and Uncertain Structural Specifications in a Linear Simultaneous Equation Model," unpublished Ph. D. dissertation, Indiana University.

[11] JEONG, JIN-HO (1988), "A Lower Risk Reduced Form Model of Macro Dynamics : A Monetarist Model with Unknown Conjecture." mimeo, University of Cincinnati.

[12] KAKWANI, N. C. (1975), "The k-class estimators of the reduced form coefficients in simultaneous equation models, *Australian Economic Papers*, 14, 250—260.

[13] KAKWANI, N. C. and R. H. Court (1972), "Reduced form coefficient estimation and forecasting for a simultaneous equation model, *The Australian Journal of Statistics*, 14, 143—160.

[14] KNIGHT, J. L. (1983), "The structure of reduced form estimators in Linear Simultaneous Equation Models," mimeo, Indiana University.

[15] KNIGHT, J. L. (1986), "Non-normal errors and the distribution of OLS and 2SLS structural estimators," *Econometric Theory*, 2, 75—106.

[16] LARIMORE, W, E. (1983), "Predictive inference, sufficiency, entropy and an asymptotic likelihood principle," *Biometrika*, 70, 175—181.

[17] MAASOUMI. E. (1978), "A modified Stein-like estimator for the reduced form parameters of simultaneous equations," *Econometrica*, 46, 695—703.

[18] MAASOUMI, E. (1983), "Uncertain structural models and generic reduced form estimation," discussion paper, Indiana University.

[19] MAASOUMI, E. (1984), "Reduced form estimation and prediction from uncertain structural models : a generic approach," *Journal of Econometrics*, 31, 3—29.

[20] MAASOUMI, Esfandiar and JIN-HO Jeong (1988), "A Comparison of GRF and Other Reduced Form Estimators in Simultaneous Equations Models," *Journal of Econometrics*, 37, 115—34.

[21] MAASSOUMI, E. and P. C. B. PHILLIPS (1982), "Misspecification in the gereral single equation case", mimeo, Yale University.

[22] MARIANO, R. S. (1982), "Analytical small-sample distribution theory in econometrics : the simultaneous-equations case," *International Economic Review*, 23, 503—533.

[23] NAGAR, A. L. and S. N. SAHAY (1978), "The bias and mean squared error of forecasts from partially restricted reduced forms," *Journal of Econometrics*, 7, 227—243.

[24] PARK, Soo-Bin (1982), "A forecasting property of the unrestricted, restricted and partially restricted reduced form coefficients," *Journal of Econometrics*, 19, 385—390.

[25] PHILLPIS, P. C. B. (1983), "Exact small sample theory in the simultaneous equations model," in Z. Griliches and M. D. Intriligator, *Handbook of Econometrics*, vol. 1 (North-Holland, Amsterdam), chapter 8.

[26] RAO, C. R. (1973), *Linear Statistical Inference and Its Application*, 2nd. Ed. (John Wiley, New York).

[27] RHODES, G. F. Jr. and M. D. WESTBROOK (1981), "A study of estimator densities and performance under misspecification," *Journal of Econometrics*, 16, 311−337.

[28] SARGAN, J. D. (1976a), "Econometric estimators and the Edgeworth approximation," *Econometrica*, 44, 421−448.

[29] SARGAN, J. D. (1976b), "The existence of the moments of estimated reduced form coefficients," London School of Economics, Discussion Paper No. A6, presented to the Vienameeting of the Econometric Society, 1977.

[30] ZELLNER, A. (1962), "An efficient method of estimation seemingly unrelated regressions and tests for aggregation bias," *Journal of the American Statistical Association*, 57, 348−368.