Discussion Paper

Discussion Paper No. 97-01 E

Businessmen's Expectations Are Neither Rational nor Adaptive

Marc Nerlove
Til Schuermann

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Industrial Economics and International Management Series

Businessmen's Expectations Are Neither Rational nor Adaptive

by

Marc Nerlove*)
Til Schuermann**)

*)Department of Agricutural and Resource Economics,
University of Maryland
**)AT&T Bell Laboratories

January 1997

Abstract

A framework which allows for the joint testing of the adaptive and rational expectations hypotheses is presented. We assume joint normality of expectations, realizations and variables in the information set, allowing for parsimonious interpretation of the data; conditional first moments are linear in the conditioning variables, and we can easily recover regression coefficients from them and test simple hypotheses by imposing zero restrictions on these coefficients. The nature of the data, which are responses to business surveys and are all categorical, requires simulation techniques to obtain full information maximum likelihood estimates. We use a latent variable model which allows for the construction of a simple likelihood function. However, this likelihood contains multi-(four)dimensional integrals, requiring simulators to evaluate. Simulated maximum-likelihood estimation is carried out using the Geweke-Hajivassilou-Keane (GHK) method, which is consistent and has low variance. The latter is crucial when maximizing the log-likelihood directly. Identification of the parameters is achieved by placing restrictions on the response thresholds and/or the variances. We find that we can reject *both* hypotheses.

I. Introduction

How expectations are formed is an issue of profound importance in economic theory, and models of expectation formation have a long history. Testing theories of expectation formation, however, requires that we find data on this intangible object called an expectation,. Direct observations are rare. Usually, the effects of changing expectations can only be inferred indirectly through observations of the aggregated outcomes of individual decisions. Ideally we would like micro-level data on both expectations and realizations. The quarterly business surveys, such as the one conducted by the Confederation of British Industries of UK manufacturing firms or the Konjunkturforschungstelle of Switzerland, contain questions about both expectations and realizations of such variables as demand, prices, and production.

The importance of using surveys to test empirically models of expectation formation is succinctly expressed by Pesaran (1987, p. 207): "Only when direct observations on expectations are available is it possible to satisfactorily compare and contrast alternative models of expectations formation." But not all the difficulties associated with indirect testing disappear and several new ones emerge. The principal new difficulty is the categorical nature of almost all the data available. Our primary focus in this paper is on how to analyze qualitative data on expectations.

In business surveys, firms are asked questions about observed changes in their demand, production and prices, their expectations of future changes in these variables, and other aspects of the firm's behavior. Their responses are primarily ordered and categorical; that is, they answer "increase", "remains the same", or "decrease" in comparison with the previous month or quarter. As a result, standard time series techniques applied directly to the categorical data are not appropriate for testing expectations hypotheses.

Traditionally, business survey data have been analyzed by means of conditional log-linear probability models (CLLP) (Nerlove, 1983). CLLP models permit reduction of the parameter space to manageable size, but they essentially treat the data as truly discrete and unordered and are perhaps best suited for data analysis. Nerlove (1988) proposed an alternative approach of treating the survey responses as being triggered by continuous latent structural variables as they cross certain thresholds.[1] The data are arranged in a J^Q contingency table, where J is the number of categories (for the business survey model, J=3) and Q the number of variables under consideration. A standard tool in the econometrician's kit for

[1] Nerlove (1988) compares the latent variable regression to a regression of a continuous variable y on two continuous variables x_1 and x_2 and uses this to illustrate the use of correlations for determining the structural relationship among latent variables under the assumption of multivariate normality.

contingency table analysis is the method of minimum chi-square.[2] While this method may be appealing on grounds of familiarity, maximum likelihood is preferable in the present application.[3] The relationships among the latent variables can be summarized in a covariance matrix that can theoretically be estimated by maximum likelihood. However, standard ML procedures are not feasible even in small models due to problems involving the computation of multi-dimensional integrals. This is not to say that such models have not been estimated. However, conventional econometric techniques, either the pairwise calculation of polychoric correlations (Pearson and Pearson, 1922, and Olsson, 1979) by maximum likelihood (Poon and Lee, 1987) or the two-step method (Martinson and Hamdan, 1971) ignore the true multivariate nature of the data and thus bias the usual tests. The second method, although simpler, has the disadvantage of not necessarily producing a matrix of estimated correlations that is positive definite.[4]

In this paper we formulate a method for testing jointly the rational and adaptive expectations hypotheses using business survey data from two countries: Switzerland and the United Kingdom. We use data on *demand* in the form of responses about incoming orders rather than data on prices. Since the majority of firms surveyed are in manufacturing, it is unlikely that they operate in anything near to a perfectly competitive market. Thus these firms are likely to be price setters rather than price takers. Demand is less under their direct control and therefore less endogenous to their own actions than are prices.

This paper extends Horvath, Nerlove and Willson (1992), who test (and reject) only rationality of British manufacturing firms' expectations for several periods using numerical (*not* simulation) full-information maximum-likelihood (FIML) methods; it builds on Nerlove and Schuermann's earlier paper (1995). Horvath, Nerlove and Willson did not provide a specific alternative to rational expectations, as we do here. We also formulate a lower dimensional version of the rational expectations hypothesis test, which can in fact be computed numerically, and compare the results with the higher dimensional version which requires simulation-based estimation. The lower (3) dimensional version of the model eliminates the future value of expectations, and thus the effects of overlapping information sets discussed below. Since our finding that future expectations are correlated with current ones is one of the most significant of our results, with far reaching implications for all work on testing models of expectation formation, we emphasize the results based on the 4-dimensional model, the estimation of which requires a more elaborate technique.

---

[2] See Rao (1955).

[3] For a discussion of first-order asymptotic equivalence of minimum chi-square and maximum likelihood, see Rao (1961, 1963). For a direct comparison of these two methods in a simulation context, see Schuermann (1993).

[4] See Nerlove, Ross and Willson (1991).

Simulation techniques, developed by McFadden (1989), Pakes and Pollard (1989) and Hajivassiliou et al. (1990), lend themselves naturally to the estimation of latent-variable models. In these seminal papers, it is shown how various simulators can be used to calculate multivariate integrals in the context of limited dependent variable (LDV) models. In this paper, we employ the smooth recursive conditioning simulator of Geweke, Hajivassiliou and Keane (GHK) to obtain simulated maximum likelihood (SML) as a way of testing the joint hypotheses. The GHK has low variance which is crucial when maximizing the likelihood function directly.[5] A recent related paper by Pesaran and Samiei (1995) examines limited dependent variable rational expectations models using simulation-based estimators similar to those proposed here.

The latent variable framework as proposed originally by Nerlove (1988) is introduced in Section II. The rational and adaptive expectations hypotheses are formulated in this framework and are presented in Section III. We then present formulation for joint testing.

In Section IV, we discuss maximum likelihood estimation for contingency tables where we will demonstrate the infeasibility of numerical FIML procedures and show how simulation methods may be used.

In Section V, we test the joint hypothesis formulation with Swiss monthly and UK quarterly surveys of manufacturing firms.

Section VI concludes with some remarks about the general implications of our results for studies of expectation formation.

II. A Latent Variable Framework

When building models using data arranged in contingency tables, it is often useful to think of the categorical variables as being generated by underlying continuous *latent* variables. Specifically, the business survey model is part of a general class of latent variable models with an observation rule

$$y = \tau(y^*)$$

where y^* is the unobserved latent random variable, and $\tau(\ )$ is the many-to-one mapping from y^* to the discrete observed variable y. In the case of business surveys we have $\tau : \Re \rightarrow \{1,2,3\}$. In other words, $\tau(\ )$ maps the entire real line into the integer set $\{1,2,3\}$. Another better known example is the binary response model

[5] We do not describe the simulators themselves. See Hajivassiliou and Ruud (1994).

(seen in probit and logit models) where $\tau(\ )$ is just the indicator function. The model appropriate to business surveys differs from the standard LDV model in that both the left *and* right hand side observations are the result of the mapping $\tau(\ )$.

Let $z^* = (y^*_t, x^*_{1t}, x^*_{2t}, \ldots, x^*_{kt})$ be a ((k+1)x1) vector of latent dependent and independent variables that satisfies the following linear relationship

$$y^*_t = x^*_t \beta + \varepsilon_t, \tag{1}$$

where β is a $(k \times 1)$ vector of coefficients, ε_t is a disturbance term and t is a time subscript. It is assumed that $E(\varepsilon_t) = 0$ and that ε_t and x' are uncorrelated. For firm i we observe categorical indicators[6] $z_{it} = (y_{it}, x_{it})$ of the unobservable latent variables $z^*_t = (y^*_t, x^*_t)$ such that

$$z_{ijt} = \begin{cases} 1 \text{ if } z^*_{jt} \leq a_{1jt} \\ 2 \text{ if } a_{1jt} < z^*_{jt} \leq a_{2jt} \\ 3 \text{ if } a_{2jt} < z^*_{jt} \end{cases} \tag{2}$$

where j=1,...,k,k+1.[7]

We assume further that y^*_t and x^*_t are jointly normally distributed with covariance matrix. It is in general not possible to identify all elements of $\theta = (\Sigma\ ; a_{ijt}, i = 1, 2, j = 1,\ldots, k, k+1)$ separately from a single cross-section of data. In particular, consider the contingency table obtained from the bivariate latent variable distribution $h(y^*,x^*)$ with thresholds $\{a_{y1}, a_{y2}, a_{x1}, a_{x2}\}$. This table will be identical to the one generated by the distribution $h(y^*/c_1, x^*/c_2)$ and thresholds $\{ a_{y1}/c_1, a_{y2}/c_1, a_{x1}/c_2, a_{x2}/c_2\}$, where c_1 and c_2 are arbitrary constants.[8] Therefore, we may normalize each z^* to have arbitrary location and variance.[9] One common identifying restriction is to let z^*_i have unit variances; then Σ is simply a matrix of (k)(k+1)/2 correlations.

---

[6] The conditions under which the categorical survey responses for expectations and realizations can be considered as independent draws from an aggregate distribution are developed in Theil (1952). In addition to cross-sectional independence, the major requirement is that individual firm's reporting thresholds are identical.

[7] We do not consider the case of time-varying thresholds.

[8] The same logic holds with respect to a non-zero mean of the latent variables. The contingency table obtained from $h(y^*,x^*)$ and thresholds $\{a\}$ will be identical to that obtained from $h(y^*-\mu_y, x^*-\mu_x)$ and thresholds $\{a-\mu\}$. (See also Horvath, Nerlove and Willson (1992).)

[9] Specific assumptions of thresholds and variances will be treated in section II.C. In some instances, we can relax the unit variance assumption if we then restrict the thresholds to be equal.

Since the joint distribution of y and **x**, f(y,**x**), is normal, so is the conditional distribution of y given **x**, f(y | **x**). The parameter vector in (1) can be inferred from Σ using

$$E(y_t^* | x_t^*) = \beta x_t^* \tag{3}$$

$$\varepsilon_t = y_t^* - E(y_t^* | x_t^*) .$$

The estimated parameter β has the form

$$\beta = \Sigma_{y_t^* x_t^*} \Sigma_{x_t^* x_t^*}^{-1} \tag{4}$$

Maximum-likelihood estimates of θ may be obtained using theory for the estimation of polychoric correlation coefficients (Olsson, 1979), and estimates for β follow directly from (4) after replacing population with sample correlations. Preferably we wold like to estimate θ via full information maximum likelihood, but this is not feasible for (k+1) > 3 latent variables due to problems associated with the calculation of multi (k+1) - dimensional integrals. We have two formulations of the rational expectations hypothesis: one which involves three variables, in which numerical integration is possible, and the other which involves four variables, in which case we are forced to simulate.

The two-step estimator, developed by Martinson and Hamdan (1979), estimates the thresholds first from the marginal frequencies by simply inverting the univariate standard normal c.d.f. The second step calculates the correlations pairwise by iterating to a root of the sample score, conditional on the first stage estimated thresholds.[10] In achieving computational feasibility, this method ignores the true multivariate nature of the data generating process.

---

[10] Nerlove, Ross and Willson (1991) point out that the information matrix, evaluated at , provides a biased estimate of the variance-covariance matrix of the estimator because it ignores the fact that the thresholds are estimated in the first stage. Olsson (1979) presents Monte Carlo evidence suggesting that the bias is not substantial.

III. Rational and Adaptive Expectations

A. Expectation formation

We review briefly the rational and adaptive expectations hypotheses and introduce some notation before formulating the joint testing procedure.[11] Let $y^e_{it} \equiv E(y_{it}|\Omega_{it-1})$ be the expectation formed at time t-1 by firm i of variable y at time t conditional on the information set Ω_{it-1}.[12] Therefore, $y^e_{it+1} \equiv E(y_{it+1}|\Omega_{it})$ is the expectation formed at time t+1 by firm i conditional on the informtation set Ω_{it}. One problem is immediately apparent: The world is not really discrete (at least not with respect to quarterly periods of observation). Some news which occurs at the end of period t-1 and the beginning of t will affect not only y_{it} and y^e_{it} but also y_{it+1} and y^e_{it+1}. Thus, rational expectations in the current period can be expected to be correlated with expectations in the next period.[13]

Standard rationality tests usually consider one particular agent's expectations using a time series of observations $\{y_{it}, y^e_{it}, t = 1,..., T\}$. These tests can be suitably modified for a time series of cross-sections, or, as the limitations of the data dictate, to handle serial tests based on an aggregate across firms at each particular time. We test the different hypotheses directly using cross-section techniques.

The essential assumption for rational expectations is that prediction errors are uncorrelated with anything in the information set Ω_{t-1}. The regression equation commonly estimated is of the form

$$y_t = \alpha_0 + \alpha_1 y^e_t + \alpha_2 z_{t-1} + \varepsilon_t , \qquad (5)$$

where z_{t-1} could be the lagged realization, y_{t-1}, or anything else in the information set Ω_{t-1}.[14] Rational expectations requires that the prediction error ε_t be orthogonal to the entire information set (such as z_{t-1}). If the prediction error is indeed correlated with any variables in Ω_{t-1}, it implies that the forecaster has not used all available information. The problem with ovelapping information sets created by artificial periods imposed on essentially continuous data is apparent here. y^e_t depends on Ω_{t-1} but, if some news spills over into next period, y^e_t will not necessarily be

---

[11] Unless otherwise specified, we will write the relationships in terms of the continuous latent variables, not their discrete realizations.

[12] Given cross-sectional independence, $E(y_{it}|\Omega_{it-1}) = E(y_{it}|\Omega_{t-1})$, where Ω_{t-1} = .

[13] In a very important paper, Sims (1971) shows that basing a distributed lag on discrete data when the world is in fact continuous, or at least characterized by a much finer time grid, can easily lead to *forward* distributions of lag, which is, of course, the essence of the problem noted in the text.

[14] For a survey of empirical tests of the rational expectations hypothesis, see Lovell (1986).

uncorrelated with z_{t-1}. The usual test for efficiency based on equation (5) with the following parameter restrictions

$$H_0: \alpha_0 = \alpha_2 = 0, \ \alpha_1 = 1,$$

is thus problematic. Alternatively, one may test the simpler hypothesis of unbiasedness

$$y_t = \alpha_0 + \alpha_1 y^e_t + \varepsilon_t, \ \ H_0: \ \alpha_0 = 0, \alpha_1 = 1, \tag{6}$$

which is not subject to this problem. However, such a test is not very powerful either. By unbiasedness, we simply mean that agents do not make systematic errors, which implies that forecast errors should have zero mean.

Under the null of rational expectations, one implication of (5) and (6) is that $var(y_t) = var(y^e_t) + var(\varepsilon_t)$ and hence $var(y_t) \geq var(y^e_t)$. In general for business surveys, the variance of realizations has been observed to be greater than for expectations.

The Adaptive Expectations model is a much simpler model of expectation formation than the rational expectations model, although, of course, adaptive expectations can be rational under certain circumstances.[15] Agents form their expectations about tomorrow by looking at the mistake they made yesterday. They revise their expectations upward or downward based on their most recent error.

$$y^e_{t+1} - y^e_t = \gamma[y_t - y^e_t] + u_t \ , \tag{7}$$

where $0 \leq \gamma < 1$ is the adjustment factor.

Estimation of the adaptive expectations hypothesis is often done by manipulating equation (7) to arrive at an expression which contains only directly observable realizations. The parameter γ then depends on changes in realizations. It is the coefficient of past realizations in an equation explaining current realizations. The residual u_t may include factors that affect short-term expectations. This is consistent with the early formulation of the adaptive expectations hypothesis in which the agent is presumed to form expectations of long term "normal" levels of certain variables (such as prices, for example).

As shown in Muth (1960, 1961), the adaptive expectations in (7) are in fact fully rational if

---

[15] See, for example, Muth (1960). Muth's formulation is generalized in Nerlove(1967) and Nerlove, Grether and Carvalho (1979; 1995, pp. 73-76).

$$y_t = \varepsilon_t + \gamma \sum_{i=1}^{\infty} \varepsilon_{t-i}$$

so that

$$\Delta y_t = \varepsilon_t + (\gamma - 1)\varepsilon_{t-1}.$$

That is, if y_t is generated by a moving average of i.i.d. random variables, then the agent's best forecast for y_{t+1} depends only on y_t.

B. Testing for rational expectations

Pesaran (1988, Chapters 7-8, pp. 162-244) gives an exhaustive account of the problems of testing the REH and attempts to do so using direct observations on expectations. When expectations are not directly observed, the problem of testing the REH is greatly complicated by the fact that only indirect tests are possible. As Pesaran (1988, pp. 179-180) puts the matter: "From an hypothesis testing viewpoint, the REH is best characterized as a peculiar type of cross-equation restriction, relating reduced form parameters of the RE equation(s) to the parameters of the equation(s) that generate the forcing variables....This characterization underlies most *indirect* tests of the RE and suggests that one possible method of testing the REH is to see whether the cross-equation restrictions implied by the hypothesis are valid." And further (p.181): "It is important to bear in mind that the validity of cross-equation tests as tests of the REH crucially depends on the validity of the unrestricted model within which the REH is embodied. In general, the rejection of the cross-equation restrictions does not necessarily lead to to the rejection of of the REH. It can always be argued that the cross equation restrictions have been rejected not because the REH is false, but due to mis-specification in the underlying economic model." The problem of indirect testing of REH models is further complicated when future expectations are included, as is the case in recent studies of foreign exchange rate determination, or when laggeged values of the dependent variable are included, as in partial adjustment models, the problems of estimating and testing RE models are greatly complicated by serial correlation which may be induced in the reduced form disturbances of such models and identification difficulties which may be severe. (See Pesaran, 1988, pp. 183-203.)

Thus, as Pesaran (p. 207) observes, "Only when direct observations on expectations are available is it possible to satisfactorily compare and contrast alternative models of expectations formation." But not all the difficulties associaed with indirect testing disappear and several new ones emerge. The principal new difficulty is the categorical nature of most of the data available. This is our primary focus in this paper. Further complications, similar to those found in dynamic behavioral models with future expectations, arise: a major source of serial correlation among future and current expectations is overlapping information sets due to temporal aggregation, as pointed out above.[16] We regard this as the most significant problem we have encountered in this investigation and wish to emphasize that it is likely to occur in general in all attempts to assess models of expectations formation. An additional problem is the issue of the underlying normality of the latent variables and stochastic disturbance. An assumption of normality is crucial to our analysis and we do not relax it here.[17]

We have data on four series from sequential pairs of surveys: y_t, y_{t-1}, y^e_{t+1}, and y^e_t, any of which may contain information about the others. We can put all of these variables together to test jointly the hypotheses of adaptive and rational expectations subject to the caveats and limitations noted above. It is important to keep in mind exactly how the data are presented to us. All we have is a $3^{(k+1)}$ contingency table filled with observed frequencies. If one casts the model into a latent variable framework as was done in Section II, one can recover estimates of model parameters by assuming that all variables are jointly normally distributed and then use properties of the multivariate normal distribution to find conditional means. All that is necessary in order to construct these model parameters is an estimate of the covariance matrix. This matrix is then partitioned to recover the model parameters (see equation (4)). In section IV.A. we show how maximum likelihood can be used to estimate this covariance matrix. If we are to construct any tests of hypotheses about expectation formation, we must write them in terms of these variances and covariances, the elements of the covariance matrix Σ. What are the implied moment restrictions of the rational and adaptive expectations hypotheses?

We assume that the four variables are jointly normally distributed with mean zero and covariance matrix Σ.

$$\begin{bmatrix} y^1 \\ y^2 \\ y^3 \\ y^4 \end{bmatrix} = \begin{bmatrix} y_t \\ y_{t-1} \\ y^e_{t+1} \\ y^e_t \end{bmatrix} \sim N(0,\Sigma), \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ & & \sigma_3^2 & \sigma_{34} \\ & & & \sigma_4^2 \end{bmatrix}. \tag{8}$$

Recall the problems of identification described in Section II. For a particular variable in a multivariate latent variable model, one cannot separately identify both the thresholds and the variance. Either one restricts both the thresholds to be the

---

[16] In his review of tests of the REH using direct observations on expectations, Pesaran (1988, p. 243) points out that "... since under the REH the expectations errors ... could also be correlated with the future values of the variables in the information set, the standard statistics used in testing the orthogonality property may not be valid." Our analysis reported below is subject to this caveat.

[17] Pesaran and Samiei (1995) deal with models in which expectational and other variables may be limited, which is a form of relaxation of strict normality, and devise computationally feasible techniques for the estimation of such models along line similar to those developed here.

same (in absolute value) and therefore identifies the variance, or the variance is restricted (to unity) and the thresholds are free parameters.

Our identifying restrictions are very similar to those made in Horvath, Nerlove and Willson (1992), although the problem here is more complex. Expectations and realizations (y_t^e and y_t respectively) of a variable are assumed to have the same threshold but different variances.[18] Otherwise we restrict the variance to be one and allow the thresholds to vary. At the individual firm level, equivalence of thresholds for expectations and realizations means that firms use the same "yardstick" when evaluating expected future movements in y_t as well as past realizations.[19] The covariance matrix Σ is somewhat simplified to

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & 1 & \sigma_{23} & \sigma_{24} \\ & & 1 & \sigma_{34} \\ & & & \sigma_4^2 \end{bmatrix} \tag{9}$$

The test of the rational expectations hypothesis (REH) which we propose is straight forward in the absence of complications resulting from overlapping information sets. Since the joint distribution of the four variables is assumed to be normal, so is their conditional distribution. We are interested in the following conditional mean

$$E(y_t \mid y_t^e, y_{t-1}) = \alpha_1 y_t^e + \alpha_2 y_{t-1}. \tag{10}$$

The null hypothesis to be tested is

$$H_0: \ \alpha_1 = 1, \ \alpha_2 = 0,$$

A maintained hypothesis is that α_0, the intercept term, is equal to zero and is required for identification of the model in latent variable form.

Expression (10) is a test of the REH based on a three-variable formulation of the joint density of y_t, y_{t-1}, and y_t^e. However, our assumption of the joint normality of the *four* variables y_t, y_{t-1}, y_{t+1}^e, and y_t^e implies linearity of the conditional expectation.

$$E(y_t \mid y_t^e, y_{t-1}, y_{t+1}^e) = \alpha_1 y_t^e + \alpha_2 y_{t-1}. + \alpha_3 y_{t+1}^e \tag{11}$$

---

[18] For an alternative discussion of the threshold problem see Pesaran (1988, pp. 212-221).

[19] The variance of the latent variable underlying the conditional expectations will be less than or equal to that of the realization under rational expectations. It is therefore desirable to have those variances be free parameters in the estimation.

Thus, were it not for the problem of overlapping information sets referred to above, an alternative test of the REH would be the *joint* hypothesis

$$H_0: \alpha_1 = 1, \alpha_2 = 0, \alpha_3 = 0.$$

The completely unrestricted model (11) is more general than the usual alternative to the REH, and it has the advantage, as we shall see, that it nests the adaptive expectations hypothesis as well. The problem is that in realistic situations, even when the REH is true for the underlying data, temporal aggregation and the resulting overlap of information sets makes it inevitable that $\alpha_3 \neq 0$. Below we test whether this is in fact the case. We find strong evidence that it is, which points up the generality of the problem in all studies of expectation formation. Then assuming $\alpha_3 \neq 0$, we test

$$H_0: \alpha_1 = 1, \alpha_2 = 0$$

in the *three* variable formulation. Our finding that this and, indeed, all restricted models are rejected in favor of the completely unrestricted model simply means that all simple models of expectation formation are misspecified. The existence of overlapping information sets in the measured data thus precludes a definitive finding against the REH. We think this conclusion is generally valid.

What are the implied covariance restrictions for the three and four variable versions of the rational expectations hypothesis, in addition to the identifying restrictions made previoulsy?

1. Three-variable testing without simulation

The common test for efficiency in rational expectations involves conditioning today's realization on the expectation made yesterday about today and yesterday's realization. Specifically

$$E(y_t \mid y_t^e, y_{t-1}) = \alpha_1 y_t^e + \alpha_2 y_{t-1}. \qquad (10)$$

with the restrictions for the null hypothesis as

$$H_0: \alpha_1 = 1, \alpha_2 = 0.$$

Without considering the identifying restrictions made in (9) for the *four* variable case, let us examine the coefficient experssions and determine what identifying restrictions we have to make for the *three* variable case. Following the numbering convention in (8) we can rewrite (10) as

$$E(y_1 \mid y_4, y_2) = \alpha_1 y_4 + \alpha_2 y_2 \qquad (10')$$

which yields

$$\alpha_1 = \left[\frac{\sigma_1^2}{\sigma_4^2}\right]\left[\frac{\sigma_{14} - \sigma_{12}\sigma_{24}}{1-\sigma_{24}^2}\right], \quad \alpha_2 = \left[\frac{\sigma_1^2}{\sigma_2^2}\right]\left[\frac{\sigma_{12} - \sigma_{14}\sigma_{24}}{1-\sigma_{24}^2}\right]$$

Under the null, $\alpha_2 = 0$. This yields the following covariance restriction: $\sigma_{12} = \sigma_{14}\sigma_{24}$. To test if $\alpha_1 = 1$ we need to make the additional assumption that $\sigma_{y_t}^2 = \sigma_1^2 = 1$ which is different from the identifying assumption made in (9). Because we impose a unit variance on y_t, we may leave the thresholds to be freely estimated. However, we are forced to impose the restriction of equal thresholds on y_{t-1} whose variance is estimated freely.

The joint covariance matrix to be estimated is:

$$\Sigma_{REH} = \begin{bmatrix} 1 & \sigma_{12} & \sigma_{14} \\ & \sigma_2^2 & \sigma_{24} \\ & & \sigma_4^2 \end{bmatrix}$$

with the restriction under the null being $\sigma_{12} = \sigma_{14}\sigma_{24}$.[20]

2. Four-variable version with simulation

We proceed from the less to the more restrictive model. The less restrictive alternative, as described by (11) is

$$E(y_t \mid y_t^e, y_{t-1}, y_{t+1}^e) = \alpha_1 y_t^e + \alpha_2 y_{t-1} + \alpha_3 y_{t+1}^e$$

with the maintained restriction of $\alpha_3 = 0$. The other parameters are freely estimated. As for the three variable version, we write

$$E(y_1 | y_2, y_3, y_4) = \begin{bmatrix} 1 & \sigma_{23} & \sigma_{24} \\ & 1 & \sigma_{34} \\ & & \sigma_4^2 \end{bmatrix}\begin{bmatrix} \sigma_{12} \\ \sigma_{13} \\ \sigma_{14} \end{bmatrix} = \begin{bmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_1 \end{bmatrix}.$$

The implied covariance restrictions are

$$\sigma_{13} = \sigma_{23} = \sigma_{24} = \sigma_{34} = 0.$$

Thus we estimate the following covariance matrix:

[20] This restriction is analogous to the one imposed by Horvath, Nerlove and Willson (1992).

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & 0 & \sigma_{14} \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & \sigma_4^1 \end{bmatrix}.$$

For the null of rational expectations, we consider the conditional expectation in (10) which is the same as the three variable case. However, to cast it fully into the joint *four* variable framework, we simply make the following additional covariance restrictions:

$$\sigma_{12} = 0,\ \sigma_{14} = \sigma_4^2 \equiv \bar{\sigma}.$$

Thus the covariance matrix for the *four* variable version of the null of rational expectations is

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & \bar{\sigma} \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & \bar{\sigma} \end{bmatrix}.$$

C. Testing for adaptive expectations

The test of the adaptive expectations hypothesis (AEH) is almost as simple. Recall that we are interested in

$$y_{t+1}^e - y_t^e = \gamma[y_t - y_t^e] + u_t$$

which is equation (7). We need to assume that u_t is uncorrelated with y_t and y_t^e. Again, taking into account the multivariate setup, the conditional expectation of interest is

$$E(y_{t+1}^e | y_t^e, y_t) = \beta_1 y_t^e + \beta y_t. \tag{12}$$

We test the restriction:

$$\beta_1 + \beta_2 = 1,$$

where $\beta_1 = (1\text{-}\gamma)$ and $\beta_2 = \gamma$, presuming $0 < \gamma < 1$. We call this the strong form of the AEH. Again, the completely unrestricted model contains an additional variable: y_{t-1}. The appropriate alternative against which to test the AEH is given by

$$E(y_{t+1}^e | y_t^e, y_t, y_{t-1}) = \beta_1 y_t^e + \beta_2 y_t + \beta_3 y_{t-1}, \tag{13}$$

where the restriction $\beta_3 = 0$ is imposed. We call this the weak form of AEH where future expectations are related to past expectations and current realizations (but *not* lagged realizations). The strong form puts a particular restriction on this relation, namely that $\beta_1 + \beta_2 = 1$. In Section V we first test the weak form against the completely unrestricted model and then proceed to the test of the strong against the weak form.

What are the implied covariance restrictions? For the AEH we remain with the four variable joint framwork throughout. Again, we proceed from the less to the more restrictive model. The AEH alternative, or its weak form, is described by (13) with $\beta_3 = 0$ being the only restriction. As with the rational expectations hypothesis, we will follow the numbering convention of (8) and write

$$E(y_3|y_1,y_2,y_4) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{14} \\ & 1 & \sigma_{24} \\ & & \sigma_4^2 \end{bmatrix} \begin{bmatrix} \sigma_{13} \\ \sigma_{23} \\ \sigma_{43} \end{bmatrix} = \begin{bmatrix} \beta_2 \\ \beta_3 \\ \beta_1 \end{bmatrix}.$$

The implied covariance restrictions are

$$\sigma_{12} = \sigma_{23} = \sigma_{24} = 0.$$

We therefore estimate the following covariance matrix:

$$\begin{bmatrix} \sigma_1^2 & 0 & \sigma_{13} & \sigma_{14} \\ & 1 & 0 & 0 \\ & & 1 & \sigma_{34} \\ & & & \sigma_4^2 \end{bmatrix}.$$

For the null of the strong form of adaptive expectations, we consider the conditional expectation given by (12) with the restriction that $\beta_1 + \beta_2 = 1$. The implied additional covariance restrictions are

$$\sigma_{14} = 0, \sigma_{13} = 1 - \sigma_{14}.$$

The strong form of the AEH implies the following covariance matrix:

$$\begin{bmatrix} \sigma_1^2 & 0 & (1-\sigma_{34}) & 0 \\ & 1 & 0 & 0 \\ & & 1 & \sigma_{34} \\ & & & \sigma_4^1 \end{bmatrix}.$$

Although the rational and adaptive expectations hypotheses are not nested one inside the other, in our representation they are both nested in a more general expectations formation framework, which allows for the effects of y_{t+1}^e in the joint distribution of the *four* variables y_t, y_{t-1}, y^e_{t+1}, and y^e_t. The rational and adaptive expectations hypotheses state that realizations and expectations are related in a particular way. In contrast, in more conventional analyses the alternative to the rational expectations hypothesis in isolation simply states that expectations are *not* rational; one has, in fact, no idea what they may be if the null of rational expectations is rejected. In our case, the alternative is estimated.

Any variable of this set of four variables, y_t, y_{t-1}, y^e_{t+1}, and y^e_t, may in general contain information about the others. For example, expectations may contain information about realized values. What might this information be in the case of the rational or adaptive expectations hypothesis? Do expectations (and only expectations) contain information about realizations? Do expectations and realizations (but not lagged realizations) contain information about future expectations? Our formulation allows us to answer such questions within one structure: the covariance matrix for the four variables in question.

IV. Full Information Maximum Likelihood Estimation

A. Formulation of the sample likelihood

In Section II we outlined a structural latent variable model which is useful for estimation with business surveys. Following equation (1), the data can be arranged into a $3^{(k+1)}$ contingency table, where (k+1) is the total number of variables in the model. A useful way of writing down the likelihood is to consider that there are $M = 3^{(k+1)}$ different regimes. Any given observation can fall into only one cell in the M-cell contingency table. Thus, if we let n_j be the observed frequency in the j^{th} cell (or the j^{th} regime), then the model likelihood is

$$\ln L(\theta) = C + \sum_{j=1}^{M} n_j \ln \pi_j(\theta),$$

where C is a constant of proportionality, and the parameter vector θ contains all thresholds and covariances. For the purposes of this paper, (k+1) = 4, and therefore we have M = 81 regimes.

Let us consider the form of the log-likelihood for the simple bivariate case. The data will fall into a 3x3 contingency table. Let n_{rs} be the observed frequency associated with the rs^{th} cell, and let π_{rs} be the probability that an observation falls into cell rs.[21] The log-likelihood is

---

[21] Thus, $\pi_{rs} = \Pr\{y=r, x=s\} = E(n_{rs})$.

$$\ln L(\theta) \propto n_{11} \ln\pi_{11} + n_{12} \ln\pi_{12} + n_{13} \ln\pi_{13}$$
$$+ n_{21} \ln\pi_{21} + n_{22} \ln\pi_{22} + n_{23} \ln\pi_{23} \qquad (14)$$
$$+ n_{31} \ln\pi_{31} + n_{32} \ln\pi_{32} + n_{33} \ln\pi_{33}.$$

The associated score can be written as

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \sum_{r=1}^{3} \sum_{s=1}^{3} [W_{rs}(n_{rs} - \pi_{rs})], \qquad (15)$$

where

$$W_{rs} = \frac{\partial \ln \pi_{rs}}{\partial\theta} = \frac{1}{\pi_{rs}} \frac{\partial \pi_{rs}}{\partial\theta}. \qquad (16)$$

We can think of (16) as a weighted residual, the residual $[n_{rs} - \pi_{rs}]$ being the deviation of the sample frequency from its expectation.

As Hajivassiliou, McFadden and others have pointed out, the primary obstacle to computing the FIML estimates for θ is the calculation of the probabilities π_{rs} which appear both linearly and nonlinearly in the likelihood and score expressions. Each of these probabilities involves integration over an (k+1) dimensional space. It will be instructive to consider this integration more carefully.

B. The integration problem

The computational barrier for latent variable and multinomial probit (MNP) models is reached when the degree of integration is greater than three. However, even in the case of a trivariate latent variable or a four-alternative MNP model, the computational burden is considered to be exceptionally high. To illustrate the problem of integration, consider a particular probability expression of the bivariate business survey model more closely. Specifically, consider the cell for values where (y = 2, x = 3) from equation (14). Note that an asterisk on a variable denotes the unobservable continuous latent variable rather than its observable categorical analogue (no asterisk).

$$E(n_{23} = \pi_{23} = \Pr\{y = 2, x = 3\} = \Pr\{a_{y1} < y^* \leq a_{y2}, a_{x2} < x^*\} \qquad (17)$$

Rewrite this joint probability as

$$\pi_{23} = \int_{a_{x2}}^{\infty} \int_{a_{y1}}^{a_{y2}} f(y^*, x^*) dy^* dx^*$$
$$= \iint I(a_{y1} < y^* \leq a_{y2}, a_{x2} < x^*) f(y^*, x^*) dy^* dx^*, \qquad (18)$$

where $I(\cdot)$ is the indicator function, $f(y^*, x^*)$ is the bivariate normal density, and the range of integration in the second part of (18) is the entire real line. This can be rewritten in a more general form as

$$\Pr(D(z);\mu,\Sigma) = \int_{D(z)} N(z-\mu)dz) \equiv E_z[I(z \in D(z))], \tag{19}$$

where z is a normal (k+1)-dimensional vector with mean μ and covariance matrix Σ, and $N(\cdot)$ is the multivariate normal density. $I(z \in D)$ is the indicator function defined for the event $D(z) = \{z \mid a_1 < z < a_2\}$. A leading case is the negative orthant probability which is particular to the MNP model, where the conditioning region $D(z) = \{z \mid z < 0\}$. Hajivassiliou, McFadden and Ruud (1991) outline and compare several simulators designed to estimate the probabilities π and its derivatives.

C. Simulated maximum likelihood (SML)

The idea of using simulations for the estimation in a maximum likelihood context is not new. Lerman and Manski (1981), for the case of the MNP model, attempted to maximize the likelihood function directly by simulating the probability expressions which appeared in the likelihood functions. The simulated maximum likelihood (SML) approach does not yield unbiased estimates of the parameters of interest for a fixed number of replications.[22] The reason for this is simple. Because of the concavity of the logarithm transformation, the probabilities do not enter the likelihood function expression linearly, so neither do the simulation errors of those probabilities.[23] If we denote $\tilde{f}$ as a simulator for the likelihood f, then $E\tilde{f} = f$. However, $E(\ln \tilde{f}) \neq \ln E(\tilde{f})$. Specifically,

$$B(\tilde{f}) \equiv E(\ln \tilde{f}) - \ln E(\tilde{f}) = -\frac{var(\tilde{f})}{2f^2} < 0.$$ [24]

The bias depends on the variance of $\tilde{f}$ (the simulation noise) which will be positive for any finite number of replications.[25]

---

22 Hajivassiliou and Ruud's chapter in Volume IV of the *Handbook of Econometrics* (1994) has an excellent discussion of this point.

23 This result is a straight forward application of Jensen's inequality.

24 See Börsch-Supan and Hajivassiliou (1990), p.15.

25 Consistency and asymptotic unbiasedness is achieved for a finite number of replications when
$\theta_R - \theta_0 = O(1/\sqrt{n})$,
whereas in the SML case we need R (number of replications in the simulator) and n (number of sample observations) to increase without bound, but at the rate $R/\sqrt{n}$;
$\theta_R - \theta_0 = O(\sqrt{n}/R)$.
(See also Gourieroux and Monfort (1991) for a special case of a simulated pseudo-ML estimator which is consistent for finite R).

As we will be estimating a four-dimensional model with fourteen parameters (six free thresholds, six covariances and two free variances), the complexity and number of first-order conditions is rather high. To make efficient use of MSM for the estimation of moment conditions such as (15) for a general weight matrix would require the simulation of good instruments. McFadden (1989) stresses that this weight matrix be constructed using simulations that are independent of $\pi(\theta)$. In the MNP case, he suggests using a polynomial of the regressors for an initial W and construct the ideal W only in the final iteration. We cannot do this, for we do not have regressors in the conventional sense. The computational burden for the construction of the ideal W at each iteration across θ would be very high indeed.

Hajivassiliou and McFadden (1990) have devised a method for simulating directly the score of the likelihood contribution which they call the method of simulated scores (MSS). Consistency is obtained for both MSM and MSS for a fixed number of random draws in the simulator. MSS was developed for the estimation of LDV models and is based on a suggestion by Ruud (1986) that the score for the general linear exponential model can be written as conditional expectations which can be simulated directly. Since the business survey model has no regressors in the conventional sense, the score expression can *not* be written in terms of conditional expectations. We thus settle for SML despite its poorer asymptotic properties.

Our view is that not much is sacrificed by using SML. While consistency is obtained by MSM and MSS for a fixed number of random draws or replications (R), efficiency is not. SML is inconsistent for a fixed number of replications in that a law of large numbers (applied to sample size) does not succeed in eliminating bias. Full efficiency is gained by MSM and MSS only if $R/\sqrt{n} \rightarrow \infty$. However, this requirement is the same as that for eliminating bias for SML (see also footnote XX). Moreover, if $R/\sqrt{n} \rightarrow \infty$, maximum likelihood efficiency is obtained as well, at which point SML has the same asymptotic properties as exact ML.[26]

Variance reduction

One may reduce the (small sample) bias of the SML estimator by incorporating a bias correction factor which will be a function of the variance of the simulator itself. Recall that $B(\tilde{f}) = -\mathrm{var}(\tilde{f})/2f^2$. Rewrite this as $\lambda\,\mathrm{var}(\tilde{f})$, where $\lambda = (2f^2)^{-1}$. We know that $\mathrm{var}(\tilde{f})$; it is simply the variance of the simulated probability

[26] See also Schuermann and Weeks (1994), who make the point that SML has the additional attraction of being rather easy to implement since writing down the likelihood function for most LDV models is rather straight forward. By contrast, McFadden and Ruud (1991) point out that both the Tobit model and Heckman's selection model are quite difficult to cast into the MSM or MSS frameworks, as are many other LDV models.

across R replications, and $\lambda \in (½,\infty)$. we have found the gain to be minimal for various (small) values of λ.

Another variance reduction technique called antithetic variates is much simpler and easier to implement. Virtually all simulators, and the GHK is no exception, have as their basic buidling block a uniform pseudo-random variate, usually on the unit interval. Let u_1 be a draw from a U[0,1]. Then another valid draw would be $u_2 = 1-u_1$. So instead of drawing a vector of length R from U[0,1], merely draw R/2 and construct their symmetric partners. The variance of this R-length vector of U[0,1] variates will be lower than if we had drawn all of them directly. We use antithetic variates in our own simulations.

V. Testing the expectations hypotheses with Swiss and British survey data

A. The data

Do Swiss and British manufacturing firms form their expectations adaptively, rationally, naively or in some other way? We attempt to answer that question here. Horvath, Nerlove and Willson (1992) use price realizations and expectations with data from the UK survey to test forecast rationality. In this paper, we use demand in the form of incoming orders instead. Since the majority of firms surveyed are in manufacturing, it is unlikely that many operate in anything near to a perfectly competitive market. These firms are therefore quite likely to be price setters rather than price takers. Demand may be less under their direct control and therefore less endogenous to their own actions. Of course, whether the firm has control over prices and demand is a question of degree, and, from this standpoint, demand seems more attractive. If a test is based on prices, in effect it tests whether expectations, *on the basis of which prices are set*, are rational, adaptive or neither.

The Swiss survey is conducted monthly whereas the British one is quarterly.[27] In addition, the horizon over which expectations are taken are different for the two surveys. The KOF asks firms what their expectations are for the following *three* months while the CBI asks firms to consider a *four* month horizon. To adjust for the different surveying frequencies, we used the October survey for the 4th quarter equivalent and January for the 1st quarter equivalent, etc. The UK sample consists of 1008 manufacturing firms for the 4th quarter of 1986 (t-1) and the 1st quarter of 1987 (t). The Swiss sample contains 942 firms.

Some Monte Carlo experimentation suggests that R = 20-30 replications for the simulator are sufficient, but because we may be dealing with potentially small

---

[27] The KOF also conducts a quarterly survey which asks, among other things, questions on price expectations and realizations.

probabilities which are notoriously difficult to estimate, we decided to increase the number of replications to 100. Furthermore, by employing antithetic variates we reduce the overall variance of the GHK simulator, critical when maximizing simulated log-likelihood functions.. The Appendix gives parameter estimates for the completely unconstrained model. Of principal interest to us, however, are not the point estimates themselves but rather the value of the log-likelihood at the respective constrained and unconstrained optima.

While about 1000 observations for each country seems ample at first glance, because we summarize these observations into a contingency table, the amount of information used in the final estimation is actually much smaller. In the four-dimensional case, we have an 81-cell contigency table yielding, of course, 80 independent empirical probabilities. For the case of three dimensions, we have 27 cells. The estimation centers on matching the simulated theoretical probabilities with the empirical ones in the contingency table. The success of this estimation depends critically on how well the empirical frequencies are themselves represented. In other words, how many observations underlie a given cell count? If that number is small (or worse, zero), this empirical estimate is poor and the likelihood function flattens out.

A solution is to pool the observations from the two countries. This in effect doubles the number of underlying observations in the contingency table and will give us more reliable empirical frequency estimates. One caveat is the differing time horizon over which managers are asked to form their expectations. We believe, however, that the benefit obtained by pooling far outweigh the costs of the slightly different expectations time horizons (three vs. four months).

B. The tests

Table 1 contains a summary of all results. In the *four* dimensional cases, each alternative expectations hypothesis is nested in the completely unconstrained model, and, of course, each null is in turn contained in its alternative. We reject both the rational and adaptive expections hypotheses, for the four as well as for the three dimensional cases, all with *p*-values of better than 0.001. For the four-dimensional models, we also reject the respective alternatives against the fully unrestricted model. All tests are likelihood ratio tests.

Table 1 Likelihood Ratio Tests

model (r)	*LRT*
AEH alt. vs. unrestricted (4)	83.08
AEH null vs. AEH alt. (2)	181.72
REH alt. vs. unrestricted (3)	192.12
REH null vs. REH alt. (2)	189.4
3-D REH test (1)	36.9

Our findings are consistent both with the extensive non simulation work by Horvath, Nerlove and Willson using 18 quarters of UK surveys, as well as our own earlier findings (Nerlove and Schuermann, 1995), in which we did not make the additional 3-D REH test and used non-pooled data.

VI. Summary & Conclusion

The testing of expectations hypotheses, so prevalent in economic theory and in fact at the core of modern macroeconomic theory, is a major challenge to the applied researcher. The business surveys which form the empirical basis for this study are a rich source of data for precisely the reason that they contain questions about the expectations managers of firms form about future demand and prices. And because we have data for these firms over time, we can proceed to subsequent periods and check the realizations against those expectations formed earlier.

A major problem is that these data come in the form of contingency tables. Any model and subsequent estimation method must make do with the J^Q contingency tables, where J is the number of categories (for us three), and Q the total number of variables in the model (for us, three or four). By assuming that the data is generated from a joint latent multivariate normal distribution, we are able to write down parameter restrictions which are implied by the expectations hypotheses.

For values of Q larger than three, simulation-based methods must be used to maximize the likelihood with respect to the parameters of interest. Traditionally, these models had been estimated by means of conditional log-linear probability models, but with recent advances in simulation-based inference, the computational barrier imposed by high-dimensional integratrion has been broken.

We test rational and adaptive expectations using demand data from about 1000 manufacturing firms in the United Kingdom and about 900 manufacturing firms in Switzerland, and we reject both hypotheses in favor of a more general formulation about expectations which contain future expectations. The problem is that rejection may, and probably does, rest on the discreteness of the periods over which expections and realizations are measured relative to the continuous nature of information flow and expectation formation. This is the same problem pointed out by Sims (1971) many years ago and is, in our view, general for all attempts to assess the valididity of empirically implementable models of expectation formation. Either one believes that the world is discrete with the coarseness implied by our measurements and rejects all simple models of expectation formation, such as rational and adaptive, or one holds fast a priori to a model of expectation formation and rejects the data! Perhaps a useful intermediate approach would be to investigate various hypotheses using data for which variable intervals of observation can be constructed.

The latent variable model and the ensuing tests, in both the three- and the four-variable cas, all rest on the assumption of multivariate normality. How strong is this assumption? We have performed goodness-of-fit (GFI) tests both for three as well as for four dimensions and reject them both. This clearly begs the question: is the rejection of the expectations hypotheses a result of a general model misspecification or rather a simple violation of the underlying distributional assumptions? The GFI tests may in fact have failed simply because of the lumping of survey responses, particularly for the "no change" category. In our view the issues raised by overlapping information sets constitute the more serious difficulty.

References

Börsch-Supan, Axel and Vassilis Hajivassiliou (1990), "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models," Cowles Foundation Discussion Paper No. 960.

Breslaw, Jon A. and James McIntosh (1991), "Multivariate Analysis of Ordered Categorical Data," mimeo, Concordia University.

Brown, Bryan W. (1992), "Simulation-Based Inference in Nonlinear Systems with Errors-In-Variables," mimeo, Rice University.

Brown, Bryan W. and Roberto S. Mariano (1990), "Stochastic Simulations for Inference in Nonlinear Errors-in-Variables Models," mimeo, University of Pennsylvania.

Geweke, John (1990), "Efficient Simulation from the Multivariate Normal and Student Distributions Subject to Linear Constraints," mimeo, University of Minnesota.

Gourieroux, Christian and A. Monfort (1991), "Simulation-Based Inference," INSEE Working Paper #37/G301.

Hajivassiliou, Vassilis, Daniel L. McFadden and Paul Ruud (1991), "Simulation of Multivariate Normal Orthant Probabilities: Methods and Programs," mimeo, Yale University.

Hajivassiliou, Vassilis, and Daniel L. McFadden (1990), "The Method of Simulated Scores for the Estimation of LDV Models with an Application to External Debt Crisis," Cowles Foundation Discussion Paper No. 967.

Horvath, Balazs, Marc Nerlove and Douglas Willson (1992), "A Reinterpretation of Direct Tests of Forecast Rationality Using Business Survey Data," mimeo, University of Pennsylvania.

Ivaldi, Marc (1990), "Statistical Methods and Business Surveys," ch. 4 in: Jean-Pierre Florens, Marc Ivaldi, Jean-Jacques Laffont and François Laisney (eds.), *Microeconometrics: Surveys and Applications*, Oxford, Basil Blackwell, 85-122.

Keane, Michael P. (1990), "A Computationally Practical Simulation Estimator for Panel Data, with Applications to Estimating Temporal Dependence in Employment and Wages," mimeo, University of Minnesota.

Lerman, Steven R. and Charles F. Manski (1981), "On the Use of Simulated Frequencies to Approximate Choice Probabilities," ch. 7 in: Charles F. Manski and Daniel L. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, MIT Press, 305-319.

Lovell, Michael C. (1986), "Test of the Rational Expectations Hypothesis," *American Economic Review*, **76**, 110-124.

Martinson, E.O., and M.A. Hamden (1971), "Maximum Likelihood and Some Other Asymptotically Efficient Estimators of Correlation in Two-way contingency," *Journal of Statistical Computation and Simulation*, **1**, 544-565.

McFadden, Daniel and Paul Ruud (1991), "Estimation by Simulation", mimeo, University of California at Berkeley.

McFadden, Daniel L. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, **57**, 995-1026.

Muth, J. F. (1960), "Optimal Properties of Exponentially Weighted Forecasts." *Journal of the American Statistical Association*, **55**: 299-306.

Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica*, **29**, 315-335.

Nerlove, Marc (1988), "Analysis of Business-Test Survey Data by Means of Latent Variable Models," in: W. Franz, W. Gaab and J. Walters (eds.), *Theoretische und Angewandte Wirtschaftsforschung, Heinz König zum 60.*, Heidelberg, Springer Verlag.

Nerlove, Marc (1983), "Expectations, Plans and Realizations in Theory and Practice," *Econometrica*, **51**, 1251-1279.

Nerlove, Marc (1967), "Distributed Lags and Unobserved Components in Economic Time Series," *Ten Economic Studies in the Tradition of Irving Fisher*, W. Fellner, et al. eds, New York, Wiley.

Nerlove, Marc, D. M. Grether, and J. L. Carvalho (1979, 1995), *Analysis of Economic Time Series: A Synthesis*, San Diego, Academic Press.

Nerlove, Marc and Til Schuermann (1995), "Expectations: Rare They Rational, Adaptive or Naive?: An Essay in Simulation-Based Inference", in G.S. Maddala, P.C.B. Phillips and T.N. Srinivasan (eds.), *Advances in Econometrics and Quantitative Economics*, 354-381, Cambridge, Blackwell.

Nerlove, Marc and Melvyn Weeks (1992), "The Construction of Multivariate Probability Simulators with an Application to the Multinomial Probit Model," mimeo, University of Pennsylvania.

Nerlove, Marc, David Ross and Douglas Willson (1993), "The Importance of Seasonality in Inventory Models: Evidence from Survey Data," *Journal of Econometrics*, **55 (1-2)**, 105-128.

Olsson, Ulf, (1979), "Maximum Likelihood Estimation of the Polychoric Correlation Coefficient", *Psychometrica*, **44**, 443-460.

Pakes, Ariel (1986), "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, **54**, 755-784.

Pakes, Ariel and David Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, **57**, 1027-1057.

Pearson, K. and E.S. Pearson (1922), "On Polychoric Coefficients of Correlation," *Biometrica*, **14**, 127-156.

Pesaran, M. H. (1988), *The Limits to Rational Expectations.* New York: Basil Blackwell, Inc.

Pesaran, M. H., and H. Samiei (1995), "Limited-Dependent Rational Expectations Models with Future Expectations, *Journal of Economic Dynamics and Control,* **19**, 1325-1353.

Poon, W.Y., and S.Y. Lee (1987), "Maximum Likelihood Estimation of Multivariate Polyserial and Polychoric Correlation Coefficients", *Psychometrica*, **52**, 409-430.

Rao, C.R. (1963), "Criteria of Estimation in Large Samples," *Sankhya*, **25**, 189-206.

Rao, C.R. (1961), "A Study of Large Sample Test Criteria through Properties of Efficient Estimators, *Sankhya*, **23**, 25-40.

Rao, C.R. (1955), Theory of the Method of Estimation by Minimum Chi-Square," *Bulletin de l'Institut Statistique International*, **35**, 25-32.

Ruud, Paul A. (1986), "On the Method of Simulated Moments for the Estimation of Limited Dependent Variable Models," mimeo, University of California at Berkeley.

Sims, C. (1971), "Discrete Approximations to Continuous Time Distributed Lags in Econometrics." *Econometrica*, **39**: 545-564.

Schuermann, Til (1993), "Three Essays in Simulation-Based Econometrics: Methods, Expectations and ARCH," Ph.D. Dissertation, University of Pennsylvania.

Schuermann, Til and Melvyn Weeks (1994), "Why You May Need Not Worry About Finite Sample Bias in Simulated Maximum Likelihood Estimation", University of York, Discussion papers in Economics No. 94/18.

Theil, Henri (1952), "On the Time Shape of Economic Micro-Variables and the Munich Business Test, *Revue de l'Institute Internationale de Statistique* **20** , 105-120.

Willson, Douglas (1991), "Inventory Investment: International Evidence from Business Survey Data," mimeo, Concordia University.