DECEMBRE 1983

N° 8328

DIRECT TEST OF THE RATIONAL EXPECTATION HYPOTHESIS

(with special attention to qualitative variables)

Christian GOURIEROUX*

Jacqueline PRADEL**

* Université Paris IX Dauphine et CEPREMAP.

** Université Paris I.

1. INTRODUCTION

In this paper, we examine the possibility of testing the rational expectation hypothesis from simultaneous observations of some variables y and of their expectations y^*. Such a direct approach has been considered by several authors (see the joint references).

These authors study in general the formation of inflationary expectations from the Livingstone series. The main advantage of a direct approach is that it does not require an a priori model connecting the variables to their expectations ; therefore, the tests which will be introduced are model-free (or non parametric, in the usual statistical terminology).

The test procedures are related to the kind of data which are available ; in particular, a great number of observations are obtained from tendency surveys and are qualitative.

The definition of the optimal prediction for qualitative variables and the determination of the associated tests of the R.E hypothesis constitute the major part of our contribution.

Since Muth [1961], an expectation is said to be rational "if it is the same as the prediction of the theory". This optimal prediction is usually defined as the best approximation of y based upon an "information set" $\mathcal{J}$, in the sense of the minimum mean squared error. In the sequel this optimal prediction will be denoted by $\mathcal{E}(y/\mathcal{J})$.

The question is : are the observed expectations y^* compatible with the above rationality ? To answer this, we have to test the null hypothesis :

$$H = \{\exists \mathcal{J} : y^* = \mathcal{E}(y/\mathcal{J})\}$$

...

This hypothesis and the associated test procedures will be discussed for different kind of quantitative or qualitative variables in the following sections ; more precisely, in section 2, we consider the definition of the optimal prediction $\mathcal{E}(y/\mathcal{J})$. In general, this prediction is not a conditional expectation (this explains the notation $\mathcal{E}$) : in particular, optimal prediction and conditional expectation are not the same when the variables are qualitative. This is due to the fact that y^* have to satisfy some constraints, for instance to only take the values 0 and 1.

Some testable characterizations of the hypothesis H are given in section 3.

In section 4, we examine the associated test procedures. The data are in general indexed by individual and/or by time. In the case of serial data, the usual Student's test of a=1 in the model : $y_t = a\, y_t^* + u_t$ cannot in general be used and we propose another test procedure valid for individual as for serial data.

An usual underlying assumption is the joint stationarity of the processes (y, y^*). We show that the proposed test procedure is still valid without the stationarity of the means.

...

2. OPTIMAL PREDICTIONS

2a. Definition

Let us consider a set of knowledge $\mathcal{A}$ and a problem of prediction of a K dimensional vector $y = (y_1, \dots y_K)'$. A set of information $\mathcal{I}$ is a family of random variables with values on $\mathbb{R}^K$, obtained from $\mathcal{A}$: $\mathcal{I} = \psi\ (\mathcal{A})$. The set $\mathcal{I}$ can be interpreted as the set of the possible predictions of y when the knowledge is $\mathcal{A}$.

The optimal prediction, denoted by $\mathcal{E}(y/\mathcal{I})$, is defined as a solution of the optimization problem :

$$(1) \qquad \underset{x \in \mathcal{I}}{\text{Min}} \quad E \ \| y-x \|^2 = \sum_{k=1}^{K} E\ (y_k - x_k)^2$$

Such a solution exists and is easily obtained if the information set $\mathcal{I}$ is of some particular forms : for instance if $\mathcal{I}$ is a closed vector space (in the sense of the norm $\sqrt{E\|.\|^2}$), if $\mathcal{I}$ is a closed convex set, or if $\mathcal{I}$ only contains elements which are indicator random variables. As it will be seen, these particular forms are well adapted to the different kinds of available data.

2b. Quantitative observations

The simplest case arises when the variable to be predicted is quantitative. In this case, it is usually assumed that the information set $\mathcal{I}$ is a closed vector space of square integrable random variables. The optimal prediction of y is then the orthogonal projection of y on $\mathcal{I}$. It is characterized by the orthogonality conditions :

$$(2) \qquad E\ [y - \mathcal{E}(y/\mathcal{I})]\, x = 0 \qquad \text{for all } x \in \mathcal{I}$$

Different forms may be considered for the information set and are associated with different optimal predictions. To cast some light on this point, let us assume that the variable to be predicted is indexed by time : $y = y_{t+1}$, and let us

...

denote $\mathcal{I} = \mathcal{I}_t$ the information set at t.

(i) if $\mathcal{I}_t$ is the set of all square integrable functions of present and lagged values of y : y_t, $y_{t-1}, \dots$, the information sets increase with time and the prediction of the theory is the conditional expectation given $\mathcal{I}_t$:

$$\mathcal{E}(y_{t+1}/\mathcal{I}_t) = E\ (y_{t+1}/\mathcal{I}_t)$$

(ii) If $\mathcal{I}_t$ is the set of linear combinations of $1, y_t$, $y_{t-1} \dots$, the optimal prediction $\mathcal{E}(y_{t+1}/\mathcal{I}_t)$ (1) is the linear regression of y_{t+1} on y_t, $y_{t-1} \dots$

(iii) If $\mathcal{I}_t$ is the set of linear combinations of $1, y_t, y_{t-1} \dots y_{t-p}$, the information sets do not increase with time. The information does not contain the observations with a lag strictly greater than p : the order p may be seen as a measure of the memory.

(iv) $\mathcal{I}_t$ may contain variables which are not past values of y. An interesting case occurs when $\mathcal{I}_t$ is the set of linear combinations of y_t and ${}_{t-1}y_t^*$ the prediction of y_t taken at t-1. In this case, the optimal prediction is of the following form :

$$\mathcal{E}(y_{t+1}/\mathcal{I}_t) = a\ y_t + b_{t-1}y_t^*$$

It corresponds to the adaptative scheme if:a+b = 1, a > 0, b > 0.

In the case of simultaneous observations of actual and expected values indexed by the time, we assume some regularity for the sequence of the information sets. More precisely, if $\mathcal{I}_t$ is some set of variables indexed by t, $\mathcal{I}_{t+1}$ can be deduced from $\mathcal{I}_t$ by the shift of one unit in the index of these variables ; this condition is in particular satisfied by the previous examples.

...

2c. R.E. hypothesis on the distributions [Lucas (1972)]

The variable to be predicted is often qualitative. The predictions for such a case depend upon the constraints on the answers. These answers can be probability distributions (in this section) or qualitative responses (in the next section). For instance, the respondents may not be questioned about the value of $y \in \mathbb{R}$, but about the expected probability for this variable to take a value between a_{k-1} and a_k, where the a_k, $k=0...K$, are given values :

$$a_0 = -\infty < a_1 \ldots < a_{K-1} < a_K = +\infty$$

The variables to be predicted are in fact the qualitative variables, given by :

$$(3) \qquad z_k = \begin{cases} 1 \text{ if } & y \in A_k = [a_{k-1}, a_k[\\ 0 \text{ otherwise} \end{cases}$$

for $k = 1,...,K$.

Since $\bigcup_{k=1}^{K} A_k = \mathbb{R}$, and $A_k \cap A_\ell = \emptyset$ for $k \neq \ell$, the random variables z_k's satisfy : $\sum_{k=1}^{K} z_k = 1$.

The expectations z_k^* $(k=1,...K)$ of the variables z_k $(k = 1...K)$ can be regarded as probabilities on the finite set $\{1,...K\}$:

$$(4) \quad 0 \leq z_k^* \leq 1 \text{ for } k=1,...K, \text{ and } \sum_{k=1}^{K} z_k^* = 1.$$

...

The formulation of the survey's question imposes in this case these constraints on the z_k^*'s. Therefore the information set can only contain elements satisfying these constraints . By analogy with the quantitative case, it is natural to define $\mathfrak{I}$ by

$$(5)\quad \mathfrak{I} = \{x = (x_1,\dots x_K)'; x_k \in \tilde{\mathfrak{I}}\ ;\ 0 \le x_k \le 1 \text{ for } k=1,\dots K \quad \sum_{k=1}^{K} x_k = 1\ \}$$

where $\tilde{\mathfrak{I}}$ is a closed vector space of random variables on $\mathbb{R}$, and is such that $\mathfrak{I} \neq \emptyset$. This form for $\mathfrak{I}$ simply means that the respondent has the same information set $\tilde{\mathfrak{I}}$ as for quantitative response on y, but that he must constrain his answer to satisfy (4).

Since $x_k \in \tilde{\mathfrak{I}}$ for all k, and $\sum_{k=1}^{K} x_k = 1$, we notice that $1 \in \tilde{\mathfrak{I}}$.

The information set $\tilde{\mathfrak{I}}$ is the intersection of a closed vector space and of the simplex and is in particular a closed convex set. Therefore, the minimization problem : $\underset{x \in \mathfrak{I}}{\text{Min}}\ E\|z-x\|^2$ has a solution.

$$\mathcal{E}(z/\mathfrak{I}) = (\mathcal{E}(z_1/\mathfrak{I}),\ \dots \mathcal{E}(z_K/\mathfrak{I})) \text{ which is a.s unique.}$$

2d. Qualitative prediction

In the previous example, the prediction $(z_1^*,\dots,z_K^*)$ does not take its values in the same set as the variable $z = (z_1,\dots z_K)$ which is to be predicted. This latter constraint can be imposed to the answer : this case is particularly important in practice.

In tendency surveys, for instance, the respondents are questioned about the expected direction of change for some economic variables. They have to choose a single alternative among answers

...

of the following type : "the variable is expected to go up", "the variable is expected to stay the same", "the variable is expected to go down".

As in the previous section, the variables to be predicted are qualitative ones :

$$z = (z_1, \dots z_K) \text{ with } z_k = \mathbb{1}_{A_k}(y) = \begin{cases} 1 \text{ if } y \in A_k \\ 0 \text{ otherwise} \end{cases}$$

where $\bigcup_{k=1}^{K} A_k = \mathbb{R}$ and $A_k \cap A_\ell = \phi$ for $k \neq \ell$

To the closed vector space $\tilde{\mathfrak{I}}$ representing the information set for quantitative response, we can associate the σ-field $\sigma(\tilde{\mathfrak{I}})$ generated by $\tilde{\mathfrak{I}}$. The information set $\mathfrak{I}$ will be defined by :

$$(6) \quad \mathfrak{I} = \{x=(x_1, \dots x_K) ; \; x_k = \mathbb{1}_{B_k} \text{ for } k = 1, \dots K \; ; \; B_k \in \sigma(\tilde{\mathfrak{I}}) \quad \bigcup_{k=1}^{K} B_k = \mathbb{R}, \; B_k \cap B_\ell = \phi \text{ for } k \neq \ell\}$$

Therefore, a possible answer is such that $x_k = 1$ for $k = k_0$, and $x_k = 0$ otherwise. Such an answer corresponds to the choice of the alternative k_0.

In this context, we are now discussing the existence of an optimal prediction.

Property 1

Let us denote by $P[A_k/\tilde{\mathfrak{I}}]$ the conditional probability of A_k given $\sigma(\tilde{\mathfrak{I}})$.

$x^* = (x_1^*, \dots x_K^*)'$ where

$$x_k^* = \begin{cases} 1 \text{ if } \quad P[A_k/\tilde{\mathfrak{I}}] > \max\limits_{j \neq k} \; P(A_j/\tilde{\mathfrak{I}}) \\ 0 \text{ otherwise} \end{cases}$$

is a solution of the minimization problem : $\min\limits_{x \in \mathfrak{I}} \; E\|z-x\|^2$

The optimal expectation behaviour consists in choosing the alternative which has the greatest conditional probability. It has been implicitly assumed that the maximum of the $P[A_j/\tilde{\mathfrak{J}}]$ was not attained for several different alternatives ; however, property 1 is easily modified in such a case[2].

Proof of property 1 :

We have : $E\ \|z-x\|^2 = \sum_{k=1}^{K} E\,[\mathbb{1}_{A_k} - \mathbb{1}_{B_k}]^2$

$$= 2\,[1 - \sum_{k=1}^{K} E\ \mathbb{1}_{A_k}\ \mathbb{1}_{B_k}]$$

Therefore, our problem is equivalent to the maximisation of :

$$\sum_{k=1}^{K} E\,(\mathbb{1}_{A_k}\ \mathbb{1}_{B_k}) = \sum_{k=1}^{K} E\ (E(\mathbb{1}_{A_k}\ \mathbb{1}_{B_k}/\tilde{\mathfrak{J}}))$$

$$= \sum_{k=1}^{K} E\ (\mathbb{1}_{B_k}\ P\ (A_k/\tilde{\mathfrak{J}}))$$

$$= E\ [\sum_{k=1}^{K} \mathbb{1}_{B_k}\ P(A_k/\tilde{\mathfrak{J}})]$$

It is easily seen that the variable $\sum_{k=1}^{K} \mathbb{1}_{B_k}\ P[A_k/\tilde{\mathfrak{J}}]$ is maximum if

$$B_k = \{P(A_k/\tilde{\mathfrak{J}}) > \max_{j \neq k} P(A_j/\tilde{\mathfrak{J}})\}$$

Q.E.D.

3. CHARACTERISATIONS OF HYPOTHESES H

3.a A lemma :

Let us consider the hypothesis $H = \{\exists \mathcal{J} \in I : y^* = \mathcal{E}(y/\mathcal{J})\}$ where I is a family of information sets and y^* is the observable prediction of the variable y.

Lemma : If there exists a smallest element $\mathcal{J}(y^*)$ of I, containing y^*, the three following propositions are equivalent :

(i) $\exists \mathcal{J} \in I : y^* = \mathcal{E}(y/\mathcal{J})$
(ii) $E\,||y - y^*||^2 \leq E\,||y - x||^2$ for all $x \in \mathcal{J}(y^*)$
(iii) $y^* = \mathcal{E}\,[y/\mathcal{J}(y^*)]$

Proof :

(i) $\Rightarrow$ (ii) : since $\mathcal{J}(y^*) \subset \mathcal{J}$, we have : $E\,||y - y^*||^2 = \underset{\tilde{x} \in \mathcal{J}}{\text{Min}}\, E\,||y - \tilde{x}||^2 \leq ||y - x||$ for all $x \in \mathcal{J}(y^*)$

(ii) $\Rightarrow$ (iii) : Since $E\,||y - y^*||^2 \leq \underset{x \in \mathcal{J}(y^*)}{\text{Min}} E\,||y - x||^2$

and $y^* \in \mathcal{J}(y^*)$, we deduce that : $y^* = \mathcal{E}\,[y/\mathcal{J}(y^*)]$

(iii) $\Rightarrow$ (i) : Take $\mathcal{J} = \mathcal{J}(y^*)$

Q.E.D.

3.b Quantitative response

As seen in section 2b, a natural family I_0 to be considered in the quantitative case is the family of the closed vector spaces. For expository convenience we assume that y is unidimensional :

Property 2 :

There exists $\mathcal{J} \in I_0$ such that $y^* = \mathcal{E}(y/\mathcal{J})$ if and only if : $E(y - y^*)\,y^* = 0$

Proof :

The smallest closed vector space generated by y^* is : $\mathcal{J}(y^*) = \{\lambda\, y^*,\ \lambda \in \mathbb{R}\}$, and the orthogonality condition $E(y - y^*)y^* = 0$

...

is necessary and sufficient for having $y^* = \mathcal{E}(y/\mathcal{J}(y^*))$ Q.E.D.

Therefore, the characterisation simply leads to the usual condition of orthogonality of the prediction y^* with the forecast error $y - y^*$.

3.c R.E. hypotheses on the distributions

The expectations are constrained to be probability distributions and a natural family $\tilde{I}_0$ to consider is :

$$\tilde{I}_0 = \left\{ \begin{array}{l} \mathcal{J} = (\underset{\sim}{x} = (x_1,\ldots,x_K)', \; x_k \in \tilde{\mathcal{J}}, \; 0 \leq x_k \leq 1, \; \sum_{k=1}^{K} x_k = 1) \\ \text{where } \tilde{\mathcal{J}} \in I_0 \end{array} \right\}$$

Property 3 :

If the prediction z^* of $z = (z_1,..z_K)'$ is such that ess sup $z_k^*=1$ for all $k = 1,\ldots K$, the hypothesis $\tilde{H} = \{\exists \mathcal{J} \in \tilde{I}_0 : z^* = \mathcal{E}(z/\mathcal{J})\}$ is satisfied if and only if :

$E\, z_k^* \,(z_k - z_k^*) \geq E\, z_k^* \,(z_j - z_j^*)$ for all $j, k = 1,\ldots K$.

Proof :

(i) The smallest element of I_0 containing z^* is the closed vector space generated by the components z_k^*, and then, it is the set of all linear combinations of these components.

Therefore, the elements of $\mathcal{J}(z^*)$ are of the form :

$$x = (x_1,\ldots, x_K) \text{ with } x_j = \sum_{k=1}^{K} \lambda_{jk}\, z_k^*, \; \lambda_{jk} \in \mathbb{R}.$$

The constraints $x_j \geq 0$ for all j, $\sum_{j=1}^{K} x_j = 1$, and the assumption:ess. sup $z_k^* = 1$ for all k, imply that for all j, k :

$$0 \leq \lambda_{jk} \leq 1 \text{ and } \sum_{j=1}^{K} \lambda_{jk} = 1.$$

...

Conversely, if these conditions on the λ's are satisfied,

$$x_j = \sum_{k=1}^{K} \lambda_{jk} z_k^* \geq 0 \text{ and } \sum_{j=1}^{K} x_j = \sum_{j=1}^{K} \sum_{k=1}^{K} \lambda_{jk} z_k^* = \sum_{k=1}^{K} z_k^* \sum_{j=1}^{K} \lambda_{jk}$$

$$= \sum_{k=1}^{K} z_k^* = 1.$$

Therefore, the set $\mathfrak{J}(z^*)$ is given by :

$$\left\{ \begin{array}{l} x = (x_1,\ldots,x_K) \; ; \; x_j = \sum_{k=1}^{K} \lambda_{jk} z_k^* \\ \text{with } \lambda_{jk} \geq 0 \text{ and } \sum_{j=1}^{K} \lambda_{jk} = 1 \text{ for all } j,\ k = 1\ldots K \end{array} \right\}$$

(ii) From Lemma (3.a), we deduce that the hypothesis $\tilde{H}$ is satisfied if and only if : $z^* = \mathcal{E}(z/\mathfrak{J}(z^*))$, that is if the optimization problem $\underset{x \in \mathfrak{J}(z^*)}{\text{Min}} E\,||z - x||^2$ has a solution $x = z^*$.

This problem is :

$$\underset{\lambda}{\text{Min}} \sum_{j=1}^{K} E\,(z_j - \sum_{k=1}^{K} \lambda_{jk} z_k^*)^2$$

$$\text{s.t.} \left\{ \begin{array}{ll} 0 \leq \lambda_{jk} \leq 1 \text{ for all } j,\ k = 1,\ldots,K & \\ \sum_{j=1}^{K} \lambda_{jk} = 1 & k = 1,\ldots,K \end{array} \right.$$

and is equivalent to the following one :

$$\underset{\lambda}{\text{Min}}\ S(\lambda) = \sum_{j=1}^{K} E\,(z_j - \sum_{k:k\neq j} \lambda_{jk} z_k^* - (1 - \sum_{k:k\neq j} \lambda_{kj})\, z_j^*)^2$$

s.t. : $0 \leq \lambda_{jk} \leq 1$ for all $j \neq k$

...

Since the objective function is a convex function, the first order Kühn-Tücker's conditions are necessary and sufficient. z^* corresponds to $\lambda^o = (\lambda^o_{jj} = 1, \lambda^o_{jk} = 0$ if $j \neq k)$, and is solution of the problem if and only if :

$$\left[\frac{\partial}{\partial \lambda_{jk}} S(\lambda) \right]_{\lambda = \lambda^o} \geq 0 \quad \text{for all } j \neq k = 1,\ldots K.$$

or equivalently : $E\, z^*_k (z_k - z^*_k) \geq E\, z^*_k (z_j - z^*_j)$ for all $k, j = 1,\ldots K$. Q.E.D

For instance, in the case of two alternatives, the conditions are

$$\begin{cases} E\, z^*_1 (z_1 - z^*_1) \geqslant 0 \\ E\, z^*_2 (z_2 - z^*_2) \geqslant 0 \end{cases}$$

since $z_2 - z^*_2 = z^*_1 - z_1$.

3.d Qualitative expectations

In this case, z and z^* are qualitative and the probability distribution of (z, z^*) can be summarized in a theoretical contingency table (p_{jk}), where p_{jk} is the probability that $z_j = 1$ and $z^*_k = 1$.

We consider the family $\tilde{\tilde{I}}_o$ of the information sets of the form :

$$\mathfrak{J} = \left\{ \begin{array}{l} x : (x_1,\ldots, x_K) \; ; \; x_k = \mathbb{1}_{B_k},\; B_k \in \sigma(\tilde{\mathfrak{J}}) \\ \\ \bigcup\limits_{k=1}^{K} B_k = \mathbb{R},\; B_k \cap B_j = \emptyset \text{ for } j \neq k,\; \tilde{\mathfrak{J}} \in I_o \end{array} \right\}$$

Property 4 :

There exists $\mathfrak{J} \in \tilde{\tilde{I}}_o$ such that $z^* = \mathcal{E}(z/\mathfrak{J})$ if and only if

...

for all $k = 1,\dots K$, $p_{kk} \geq \underset{j \neq k}{\text{Max}}\ p_{jk}$

Proof :

From Lemma (3.a), we know that $\tilde{\tilde{H}}$ is satisfied if and only if : $z^* = \mathcal{E}(z/\mathfrak{I}(z^*))$, where $\mathfrak{I}(z^*)$ is the smallest element of $\tilde{I}_o$ containing z^*. Using property 1, it is easily seen that the optimal prediction $\tilde{x} = \mathcal{E}(z/\mathfrak{I}(z^*))$ is such that, for all $k = 1,\dots K$ with $z_k^* = 1$:

$$\tilde{x}_j = 1 \quad \text{if } P(z_j = 1/z_k^* = 1) > \underset{j' \neq j}{\text{Max}}\ P\ (z_{j'} = 1/z_k^* = 1)$$

$$\tilde{x}_j = 0 \quad \text{if } P(z_j = 1/z_k^* = 1) < \underset{j' \neq j}{\text{Max}}\ P\ (z_{j'} = 1/z_k^* = 1)$$

(if there exist several indices $j_1 \dots j_p$ such that $P(z_{j1} = 1/z_k^* = 1) = \dots = P(z_{jp} = 1/z_k^* = 1)$, $(\tilde{x}_{j1},\dots,\tilde{x}_{jp})$ is any permutation of $(1,0\dots,0)$).

Since for all j,k: $P(z_j = 1/z_k^* = 1) = \dfrac{p_{jk}}{\sum_{j=1}^{K} p_{jk}}$ these relations can be expressed in terms of p_{jk}. Finally z_k^* is an optimal prediction if and only if :

$$\text{for all } k = 1,\dots, K \qquad p_{kk} \geq \underset{j \neq k}{\text{Max}}\ p_{jk} \qquad \text{Q.E.D.}$$

We can note that these rationality conditions are different from the conditions which are intuitively to be supposed for having more often right predictions; these conditions are :

$$P(z_k^* = 1/z_k = 1) \geq \underset{j \neq k}{\text{Max}}\ P(z_j^* = 1 / z_k = 1)$$

or equivalently : $p_{kk} \geq \underset{j \neq k}{\text{Max}}\ p_{kj}$

The concept of rationality implies the comparison of the probabilities of a same column, and the idea of right predictions implies the comparison of the probabilities of a same row.

...

3.e Testing from qualitative observations the R.E. hypothesis written for the latent quantitative variables

The qualitative variable to be predicted is often defined from a quantitative variable. In the simplest case, the possible answers are "go up" and "go down". The observable variable is :

$$z = \begin{cases} 1 & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

where y denotes the modification of the latent quantitative variable.

In the previous sections, we have considered the RE hypothesis on the qualitative variables z : $\hat{z} = \mathcal{E}(z/\mathfrak{I})$.

Another rational prediction behaviour of the respondent is to optimaly predict the latent variable by $\hat{y} : \mathcal{E}(y/\mathfrak{I})$ and to answer : $\tilde{z} = \begin{cases} 1 \text{ if } \hat{y} > 0 \\ 0 \text{ otherwise} \end{cases}$

This hypothesis can be written as :

$$\tilde{H} = \{ \exists\, \mathfrak{I} \in I_0, \; \begin{array}{l} z^* = 1 \text{ if } \hat{y} = \mathcal{E}(y/\mathfrak{I}) > 0 \\ z^* = 0 \text{ otherwise} \end{array} \}$$

where I_0 is the family of all closed vector spaces. The new difficulty is that the quantitative variables are only known through their signs and it is easily seen that the necessary and sufficient condition for the RE hypothesis to be satisfied : $E(y - \hat{y})\hat{y} = 0$ can not be written in terms of their qualitative counterparts z and $\tilde{z}$.

In such a case, it is necessary to transform in some way the qualitative data z and $\tilde{z}$ into quantitative ones.

Among the various procedures introduced in the literature two are very often used. The first one consists in replacing the observations "go up" and "go down" by scaling values. In the second one, the quantitative variable associated to a set of qualitative individual observations is the "balance" defined as the difference between the number of responses "go up" and the number of responses "go down".

...

These conversion procedures can be easily criticize , but our aim in this section is to study the implication of the R.E. hypothesis when these conversion procedures are considered to be the right ones and to compare the associated constraints on the probabilities with those obtained in the previous section 3.d.

In the sequel, we only condider the case of two alternatives. We assume that the theoretical contingency table at time t is :

z \ z^*	0	1
0	$p_{00,t}$	$p_{01,t}$
1	$p_{10,t}$	$p_{11,t}$

(i) <u>conversion by scaling technic</u> [ANDERSON (1952)] [PESARAN-GULAMINI (1982)].

We assume that there exist $\alpha < 0$ and $\beta \geq 0$ such that :

$$\begin{cases} y_{it} = \alpha \text{ if } z_{it} = 0 \\ y_{it} = \beta \text{ if } z_{it} = 1 \\ y^*_{it} = \alpha \text{ if } z^*_{it} = 0 \\ y^*_{it} = \beta \text{ if } z^*_{it} = 1 \end{cases}$$

The condition : $E(y_{it} - y^*_{it})y^*_{it} = 0$ is equivalent to :

$$p_{01,t}(\alpha - \beta)\,\beta + p_{10,t}(\beta - \alpha)\alpha = 0$$

or to : $p_{01,t}\,\beta - p_{10,t}\,\alpha = 0$

...

Therefore, the R.E. hypothesis can be written as :

$$\left\{ \frac{p_{10,t}}{p_{01,t}} \text{ independent from } t \right\}$$

and is easily tested from the observed frequencies.

Independently from the simplicity of the conversion procedure, which assumes the existence of the scaling values and their time invariance, it is interesting to remark that the expression of the R.E. hypothesis only depends upon the crossed probabilities. Contrary to the case of qualitative predictions considered in 3.d, the only important expectations from a statistical point of view are the erroneous ones.

(ii) Conversion by balance[6]

If the observed frequency table is $(\hat{p}_{ij,t})$ the balances, defined as the difference between the proportion of 1 and the proportion of 0, are :

$$\begin{cases} B_t = \hat{p}_{10,t} + \hat{p}_{11,t} - \hat{p}_{00,t} - \hat{p}_{01,t} \\ B_t^* = \hat{p}_{01,t} + \hat{p}_{11,t} - \hat{p}_{00,t} - \hat{p}_{10,t} \end{cases}$$

The R.E. hypothesis written on the quantitative variables B is : $E(B_t - B_t^*)\, B_t^* = 0$ for all t, and is equivalent to :

$$E(\hat{p}_{10,t} - \hat{p}_{01,t})\,(2(\hat{p}_{01,t} + \hat{p}_{11,t}) - 1) = 0$$

Asymptotically, when the sample size at time t tends to infinity, this condition becomes :

$$(p_{10,t} - p_{01,t})\,[2(p_{01,t} + p_{11,t}) - 1] = 0$$

or equivalently : $\left\{ \dfrac{p_{10,t}}{p_{01,t}} = 1 \quad \text{or} \quad p_{01,t} + p_{11,t} = \dfrac{1}{2} \right\}$

With a completely different conversion method, we find

...

again a condition of time invariance of $\frac{p_{01,t}}{p_{10,t}}$, as in the previous subsection.

4. TESTS OF THE HYPOTHESIS H

Once the hypothesis H has been expressed in mathematical terms, the test procedures are easily deduced. We only have to test by classical approaches either equalities for the quantitative case or inequalities for the qualitative ones. However we would like to insist on two points.

On the one hand, the data are in general indexed by individual and by time. In a first approach, they can be considered as individual data or as serial data, and the treatment of the disturbances heavily depends upon the kind of dimension which is considered.

On the other hand, for serial data, if the predictions are unbiased, the test procedures are valid without assuming the joint stationarity of the processes (y, y^*).

We will essentially discuss these two points for the quantitative case. The test procedures obtained for this case are easily transformed for treating the qualitative cases (3), by using classical general results on test of inequalities [see for instance CHANT (1974), CHERNOFF (1954), NUSCH (1964), GOURIEROUX-HOLLY-MONFORT (1982)].

4.a Test of the R.E. hypothesis from individual data

In this section, we consider panel data on y and y^* for a given date t_o : $(y_{it_o}, y^*_{it_o})$ for i = 1,..., n and we assume

that the individual observations are i.i.d. The R.E. hypothesis at the individual level is characterized by the condition :
$E(y_{it_o} - y^*_{it_o})y^*_{it_o} = 0$.

As it is well known, this conditions can be easily tested by writing the linear model :

$$y_{it} = a\, y^*_{it_o} + u_{it_o} \qquad i = 1\ldots n$$
$$Eu_{it_o} = 0,\ Vu_{it_o} = \sigma^2$$

...

and by using the classical student statistic associated with $a = 1$. This test is valid for individual data since under the null the disturbances u_{it_o} are uncorrelated with all the expectations y^*_{j,t_o}, $j = 1,..n.$

As we shall see in the next section, this latter condition is not general satisfied for serial data, and therefore, contrary to the usual practice [TURNOVSKI (1970), FRIEDMAN (1980), BROWN-MAITAL (1981), LEONARD (1982), PESARAN-GULAMINI (1982)], the student statistic can no longer be used.

4.b Test of the R.E. hypothesis for serial data under the stationarity assumption

The R.E. hypothesis can also be considered for the aggregate data(4) : $y_t = \sum_{i=1}^{n} y_{it}$, $y^*_t = \sum_{i=1}^{n} y^*_{it}$ and is :

$H = \{\forall t : E(y_t - y^*_t)y^*_t = 0\}$

If we assume the joint stationarity of the process (y_t, y^*_t), the hypothesis becomes : $H = \{E(y_t - y^*_t)y^*_t = 0\}$ and a test can be built from the empirical mean: $\frac{1}{T}\sum_{t=1}^{T} (y_t - y^*_t)y^*_t$. Under the null, the statistic $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} (y_t - y^*_t)y^*_t$ is asymptotically (in T) normal with zero mean and with an asymptotic variance given by :

$$Q = \sum_{h=-\infty}^{+\infty} \gamma(h) \qquad \text{where : } \gamma(h) = E\left[(y_t - y^*_t)y^*_t\,(y_{t+h} - y^*_{t+h})y^*_{t+h}\right]$$

This asymptotic variance is consistently estimated by replacing the infinite sum $\sum_{h=-\infty}^{+\infty}$ by a finite one $\sum_{h=-K_T}^{K_T}$ and the expectations by the empirical means [see FULLER (1976)]. Therefore, the

...

statistic ξ defined by :

$$\xi = \frac{\left[\sum_{t=1}^{T} (y_t - y_t^*) y_t^*\right]^2}{\sum_{k=-K_T}^{K_T} \frac{T}{T-|h|} \sum_{t=1}^{T-|h|} (y_t - y_t^*) y_t^* \, (y_{t+|h|} - y_{t+|h|}^*) y_{t+|h|}^*}$$

has asymptotically a $\chi^2(1)$ distribution.

If ξ is greater than the upper α point $\chi^2_\alpha(1)$ of the $\chi^2(1)$ distribution, the R.E. hypothesis is rejected. It is not rejected otherwise.

This test procedure is different from the usual one which consists in writing the linear model :

$$y_t = a\, y_t^* + u_t, \quad t = 1,\ldots,T$$

and in testing the hypothesis H by the student statistic associated with a = 1. This procedure is not in general valid in the serial case ; in effect, under the null, the disturbance $u_t = y_t - y_t^*$ is uncorrelated with y_t^* ; it may be uncorrelated with the past expectations y_{t-j}^*, $j > 0$ if the information sets are increasing with time(5). But it is in general correlated with the future expectations y_{t+j}^*, $j > 0$ therefore, the usual formula for the variance of the OLS estimator of a, and the test based on the usual student statistic are not valid.

4.c Test of the R.E. hypothesis for serial data without assuming the stationarity

The assumption of stationarity is rather strong. However, it can be replaced by a weaker assumption, if the expectations are unbiased : we assume that each of the processes y_t and y_t^* can be decomposed into a deterministic part and a zero mean random stationary component :

$$\begin{cases} y_t = f(t) + u_t \\ y_t^* = g(t) + v_t \end{cases} \qquad t = 1,\ldots T$$

...

where (u_t, v_t) is jointly stationary, $E\, u_t = E\, v_t = 0$.

(i) Let us first consider the test of unbiasedness :

$$\mathcal{H} = \{\forall t : E\, y_t = E\, y_t^*\}$$

Under $\mathcal{H}$, the process $y_t - y_t^* = u_t - v_t$ is stationary and therefore the hypothesis $\mathcal{H}$ can be easily tested from the empirical mean :

$\frac{1}{T} \sum_{t=1}^{T} (y_t - y_t^*)$ using a procedure similar to that of section 4.b.

(ii) If the hypothesis $\mathcal{H}$ is accepted after this preliminary test, we have :

$$\begin{cases} y_t = f(t) + u_t \\ y_t^* = f(t) + v_t \end{cases} \qquad t = 1,\dots,T$$

The test procedure developped in section 4.b and based on the statistic ξ is still valid in this case. This is a consequence of the equality :

$$(y_t - y_t^*)y_t^* = (u_t - v_t)\ f(t) + (u_t - v_t)v_t$$

and of the fact that the asymptotic results on the empirical means can be generalized to the product of a stationary process by a deterministic function [see Fuller (1976)].

...

FOOTNOTES

(1) The condition $1 \in \tilde{\mathcal{I}}_t$ is equivalent to the unbiasedness of the expectations : $E(y - y^*) = 0$. This condition is not necessarely to be imposed for testing rationality.

The presence of a bias simply means that the constant variables, that are the deterministic evolutions, do not belong to $\mathcal{I}_t$. These variables are of course always observables, but it can be difficult to take them into account if the transition between y_t and the constants (for instance the trend and/or the saisonnality) is unknown.

(2) In the more general case, an optimal prediction $x^* = (x_1^*, .. x_K^*)$ is defined by :

$$x_k^* = \begin{cases} 1 \text{ if } P(A_k/\tilde{\mathcal{I}}) > \underset{j \neq k}{\text{Max}}\, P(A_j/\tilde{\mathcal{I}}) \\ \\ 0 \text{ if } P(A_k/\tilde{\mathcal{I}}) < \underset{j \neq k}{\text{Max}}\, P(A_j/\tilde{\mathcal{I}}) \end{cases} \qquad \text{for } k = 1,\dots K$$

and if there exist several indices $k_1,\dots,k_p$ such that :

$$P(A_{k_1}/\tilde{\mathcal{I}}) = \dots = P(A_{k_p}/\tilde{\mathcal{I}}) > \underset{j \neq k_1,..k_p}{\text{Max}} \quad P(A_j/\tilde{\mathcal{I}}),$$

$(x_{k_1}^*, \dots, x_{k_p}^*)$ is any permutation of $(1, 0, \dots, 0)$.

(3) In the case of rational expectations on the distributions, a supplementary assumption is necessary, which cannot be tested : ess sup $z_k^* = 1$, for all $k = 1, \dots K$. Since z_k^* is observable for a particular state of nature ω_o, this assumption can not be verified from the observations. However the assumption that for some state of nature (with non zero probability) the prediction is equal to 1 is rather weak. Under this supplementary assumption, Property 3 is applicable, and leads to test the inequalities.

$$E\, z_k^* \,(z_k - z_k^*) \leq E\, z_k^*(z_j - z_j^*) \text{ for all } j \neq k$$

(4) The rationality at an aggregate level may of course be related to the rationality at an individual level.

...

Let us assume that the variable to be predicted can be decomposed into : $y_{it} = a_i + b_t$, where a_i and b_t are independent, and such that $Ea_i = 0$. Similarly we consider a decomposition of the information set : $\mathcal{I}_{it} = \mathcal{I}_i + \mathcal{I}_t$, where the individual and serial information set are independent, such that $1 \in \mathcal{I}_t$. Furthermore, we assume that $\mathcal{I}_t$ (resp. $\mathcal{I}_i$) gives no information on a_i (resp. b_t).

We have :

$$\frac{1}{n} y_t^* = \frac{1}{n} \sum_{i=1}^{n} y_{it}^* = \frac{1}{n} \sum_{i=1}^{n} \mathcal{E}(y_{it}/\mathcal{I}_{it})$$

$$= \frac{1}{n} \sum_{i=1}^{n} [\mathcal{E}(a_i/\mathcal{I}_i) + \mathcal{E}(b_t/\mathcal{I}_t)]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathcal{E}(a_i/\mathcal{I}_i) + \mathcal{E}(b_t/\mathcal{I}_t)$$

Under the assumption of independence and equidistribution of the individual components $(a_i, \mathcal{I}_i)$, we see that :

$$\lim_{n \to +\infty} \frac{1}{n} y_t^* = \mathcal{E}(b_t/\mathcal{I}_t)$$

$$\lim_{n \to +\infty} \frac{1}{n} y_t = b_t$$

and then : $$\lim_{n \to +\infty} \frac{1}{n} y_t^* = \mathcal{E}(\lim_{n \to +\infty} \frac{1}{n} y_t/\mathcal{I}_t)$$

The individual rationality implies rational expectations at the aggregate level.

(5) The R.E. hypothesis with increasing information sets may be easily tested. In effect, this hypothesis :

$$H^* = \{ \exists (\mathcal{I}_t) \subset I_o, \forall t\ \mathcal{I}_t \subset \mathcal{I}_{t+1}, Y_t^* = \mathcal{E}(y_t/\mathcal{I}_t) \}$$

...

is characterized by the conditions :

$$\forall t,\ \forall j \geq 0 \qquad E(y_t - y_t^*)y_{t-j}^* = 0$$

Let us finally remark that if $\mathcal{I}_t$ is generated by the present and past values of y, the covariance $\gamma(h), h \neq 0$ is $\gamma(h) = E[(y_{t+h} - y_{t+h}^*)\ y_{t+h}^*\ (y_t - y_t^*)y_t^*] = 0$ and the usual Student test is valid. However the assumption on the form of $\mathcal{I}_t$ cannot be rigorously tested from the data.

(6) It is not in general considered that the balances coincide with the latent aggregate quantitative variables. However, the comparison of the B_t's with the quantitative observations y_t obtained from other surveys often lead to consider that these observations are almost linearly related.

If we assume that this relation $y_t = a\ B_t + b$ is also valid for the expectation variables $y_t^* = a\ B_t^* + b$, it is easily seen that the condition $E(y_t - y_t^*)y_t^* = 0$ with the unbiasedness condition : $E(y_t - y_t^*) = 0$, is equivalent to : $E(B_t - B_t^*)B_t^* = 0$ and $E(B_t - B_t^*) = 0$ that is to the same conditions expressed in terms of B and B^*.

ANDERSON, O. (1952) : "The business test of the IFO Institute for Economic Research and its theoretical model", Revue de l'Institut International de statistique, 20, 1-17.

BROWN, B. and S., MAITAL (1981) : "What do economists know ? An empirical study of expert's expectations". Econometrica, 49, 491-504.

CARGILL, T. (1976) : "Anticipated price changes and nominal interest rates in the 1950's", Review of Economics and Statistics, 58, 364-367.

CARLSON, J. (1977) : "A study of price forecasts", Annals of Economic and Social Measurement, 6, 27-56.

CHANT, D. (1974) : "On asymptotic test of composite hypotheses in non standard conditions", Biometrika 61, 291-298.

CHERNOFF, H. (1954) : "On the distribution of the likelihood ratio", Annals of Mathematical Statistics, 25, 753-758.

FIGLEWSKI, S. and P. WACHTEL (1981) : "The formation of inflationary expectations", Review of Economic and Statistics 63, 1-10.

FRIEDMAN, B. (1980) : "Survey evidence on the rationality of interest rate expectations", Journal of Monetary Economics, 6, 453-465.

FULLER, W. (1976) : "Introduction to statistical time series", Wiley.

GOURIEROUX, C., HOLLY, A. and A. MONFORT (1982) : "Likelihood ratio test, Wald test and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters", Econometrica, 50, 63-78.

LEONARD, J. (1982) : "Wage expectations in the labour market survey evidence on rationality", Review of Economics and Statistics, 64, 157-160.

LUCAS, R. (1972) : "Expectations and the neutrality of money", Journal of Economic Theory, 4,

MULLINEAUX, D. (1978) : "On testing for rationality : another look at the Livingstone price expectations data", Journal of Political Economy 86, 329-336.

MUTH, J. (1961) : "Rational expectations and the theory of price movements", Econometrica 29, 315-335.

NÜSCH, P. (1966) : "On the problem of testing location in multivariate problems for restricted alternatives", Annals of Mathematical Statistics 37, 113-119.

PESARAN, M. and R. GULAMINI (1982) : "Expectations formations and Macroeconometric modelling", presented at the International Seminar on Recent Developments in Macroeconometric Modelling, Paris.

TURNOVSKI, S. (1970) : "Empirical evidence on the formation of price expectations", Journal of the American Statistical Association 65, 65, 1441-1454.

TURNOVSKI, S. and M. WACHTER (1972) : "A test of the expectations hypothesis using directly observed wage and price expectations", Review of Economics and Statistics, 54, 47-54.