

AN APPLICATION OF RANDOM COEFFICIENT  
REGRESSION : THE CASE OF R & D EXPENDITURE  
AND MARKET STRUCTURE.

Dennis Leech

Paul Stoneman

NUMBER 91

**WARWICK ECONOMIC RESEARCH PAPERS**

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK  
COVENTRY



AN APPLICATION OF RANDOM COEFFICIENT  
REGRESSION : THE CASE OF R & D EXPENDITURE  
AND MARKET STRUCTURE.

Dennis Leech

Paul Stoneman

NUMBER 91

July 1976

This paper is circulated for discussion purposes only and its contents should be considered preliminary.



## I. Introduction

This paper has two main aims. First, to consider the econometric problem of estimating a regression model in which the coefficients, on theoretical grounds, must be regarded as varying from one cross sectional observation to another; second, to derive, in terms of a theoretical model the exact form of the relationship between research and development expenditure and market structure. This relationship being one in which parameters vary across industries we adopt a random coefficients approach to estimate and test the average relationship. We thus develop a random coefficient regression estimator and apply it to this relationship between R & D and market structure.

The specification of the model is derived in Section II. In Section III we consider the problem raised by parameter variation within the sample. In Section IV we discuss the problem of estimating a random coefficient model and our empirical results are presented in Section V.

## II. The Model

The relationship between research and development expenditure and market structure has in the last ten years received much attention from economists. A good indication of this is the number of surveys of the area that have suddenly appeared (Grether [3], Kamien and Schwartz [6], Kennedy and Thirlwall [7]). Even so, Needham [11] can still complain of the lack of any consistent framework for analysing the relationship between R & D and market structure. Needham goes on to suggest a format based on Dorfman-Steiner conditions. In this paper we pick up this approach and in combination with an approach based on that of Cowling and Waterson [1] generate specific relationships between R & D expenditure and a measure of concentration that are amenable to estimation. This is achieved by deriving a relationship between R & D expenditure and market structure for the industry, as opposed to those derived for the firm by Needham. This level of aggregation is essentially conditioned by the fact that R & D data tends to be available only at the industry level. Moreover by explicitly detailing the derivation of this relationship we obtain a) the exact concentration measure to use and b) the form of the functional relationship between R & D expenditure and concentration.

Following Nerlove and Arrow [13], we assume that the firm maximises its present value, subject to the condition that, letting a dotted variable refer to a time derivative,

$$\dot{x}_{it}^* - x_{it} + \delta x_{it}^* = 0 \quad (i)$$

where  $x_{it}^*$  = stock of technical knowledge of firm  $i$  in time  $t$

$x_{it}$  = addition (gross) to the stock of technical knowledge of firm  $i$  in time  $t$

$\delta_i$  = rate of depreciation of technical knowledge of firm  $i$

The firm operates in an oligopolistic market with differentiated products with the inverse demand function (1)

$$P_{it} = f(Q_{1t} + Q_{2t} + \dots + Q_{nt}) + g_i(x_{1t}^*, x_{2t}^*, \dots, x_{nt}^*)$$

$$= f(Q_t) + g_i(x_{1t}^*, x_{2t}^*, \dots, x_{nt}^*)$$

where  $f_{Q_t} < 0$ ,  $g_{ii} > 0$ ,  $g_{ij} < 0$ ,

$$Q_t = \sum_{i=1}^n Q_{it}$$

$Q_{it}$  = output of firm  $i$  in time  $t$  measured in suitable units

$P_{it}$  = price of firm  $i$ 's product in time  $t$ .

- 
- (1) Our method here of introducing differentiated products follows closely that of Nickell and Metcalf [10]. It should also be noted that we are following Needham here by allowing the firms technology to affect its demand function, but we go one step further in (ii) by also allowing technology to affect costs (i.e. process innovation in addition to product innovation).

We assume that current additions to technical knowledge are related to current R & D expenditure ( $R_{it}$ ) by

$$x_{it} = x_{it} (R_{it}).$$

From the Lagrangean (ii),

$$L = \int_0^{\infty} e^{-rt} \left[ P_{it} Q_{it} - C_{it} (Q_{it}, x_{it}^*) - R_{it} \right] + \lambda(t) (x_{it}^* - x_{it} + \delta x_{it}^*) \quad (ii)$$

where  $r$  = rate of discount, we show in the Appendix that one can derive (iii)

$$R_{it} = \frac{1}{r + \delta_i} \cdot \frac{\partial x_{it}^*}{\partial R_{it}} \cdot \frac{R_{it}}{x_{it}^*} \cdot P_{it} Q_{it} \left[ \frac{\partial P_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{P_{it}} - \frac{\frac{\partial C_{it}}{\partial x_{it}^*} \frac{x_{it}^*}{C_{it}}}{\frac{\partial C_{it}}{\partial Q_{it}} \frac{Q_{it}}{C_{it}}} \right] \left[ 1 + \frac{\partial P_{it}}{\partial Q_{it}} \frac{Q_{it}}{P_{it}} \right] \quad (iii)$$

In equation (iii) we have two terms that allow for conjectural variations,

$$a) \frac{\partial P_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{P_{it}} = \left[ \frac{\partial P_{it}}{\partial x_{it}^*} + \sum_{j \neq i} \frac{\partial P_{it}}{\partial x_{jt}^*} \cdot \frac{\partial x_{jt}^*}{\partial x_{it}^*} \right] \frac{x_{it}^*}{P_{it}}$$

$$b) \frac{\partial P_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{P_{it}} = \left[ \frac{\partial P_{it}}{\partial Q_{it}} + \sum_{j \neq i} \frac{\partial P_{it}}{\partial Q_{jt}} \cdot \frac{\partial Q_{jt}}{\partial Q_{it}} \right] \frac{Q_{it}}{P_{it}}$$

By making different assumptions with respect to the conjectural variations terms  $\left( \frac{\partial x_{jt}^*}{\partial x_{it}^*} \text{ and } \frac{\partial Q_{jt}}{\partial Q_{it}} \right)$  we can derive a number of different

forms relating R & D to market structure. We are going to assume that firm  $i$  expects that (1)  $\sum_{i \neq j} \frac{\partial Q_{jt}}{\partial Q_{it}} = \lambda_i$ . (i.e. following Cowling and

Waterson [1] and Nickell and Metcalf [10]) and (11)  $\frac{\partial x_{jt}^*}{\partial x_{it}^*} = 0$  for all  $j$ , so that retaliation on R & D is taken account of by  $\delta_i$  (this is the rationale provided by Nerlove and Arrow [13] for  $\delta_i > 0$ ). (1) Given the form of the

demand function,  $\frac{\partial P_{it}}{\partial Q_{it}} = \frac{\partial P_{it}}{\partial Q_{jt}} = \frac{\partial f}{\partial Q_t}$ , and if we further assume,

using  $\eta$  to represent an elasticity, that  $\eta_{x_{it}^* R_{it}} = \eta_{x^* R}$ ,  $\eta_{C_{it} x_{it}^*} = \eta_{C x^*}$ ,

$\eta_{C_{it} Q_{it}} = \eta_{C Q}$  and  $\eta_{P_{it} x_{it}^*} = \eta_{P x^*}$  (i.e. that these elasticities

are the same for all firms in an industry in each time period), and that  $\lambda_i = \lambda$ ,  $\delta_i = \delta$ , then we can rewrite (iv) as (2)

$$R_{it} = \frac{1}{(r + \delta)} \eta_{xR} \cdot P_{it} Q_{it} \left[ \eta_{P x^*} - \frac{\eta_{C x^*}}{\eta_{C Q}} \left( 1 + [1 + \lambda] \frac{\partial f}{\partial Q_t} \cdot \frac{Q_{it}}{P_{it}} \right) \right] \quad (iv)$$

Summing over  $i$ , letting  $\sum R_{it} = R_t$ ,

(1) For a more general approach allowing for more conjectural variation see Lambin [9]

(2) These are rather drastic assumptions implying that in an industry all firms in each period are equally efficient at producing R & D results, the cost innovation/product innovation split is given, that all firms have the same constant elasticity cost functions and demand functions (w.r.t. R & D) and that expectations of retaliation are the same. We return to this in the conclusion below.

$$R_t = \frac{1}{(r+\delta)} n_{xR} \cdot n_{Px^*} \cdot \Sigma P_{it} Q_{it} - \frac{1}{(r+\delta)} n_{x^*R} \cdot \frac{n_{Cx^*}}{n_{CQ}} \cdot \Sigma P_{it} Q_{it} \\ - \frac{1}{(r+\delta)} \cdot n_{x^*R} \cdot \frac{n_{Cx^*}}{n_{CQ}} (1+\lambda) \Sigma \left[ P_{it} Q_{it} \cdot \frac{\partial f}{\partial Q_t} \cdot \frac{Q_{it}}{P_{it}} \right] \quad (v)$$

Dividing through by  $\Sigma P_{it} Q_{it}$ ,

$$\frac{R_t}{\Sigma P_{it} Q_{it}} = \frac{1}{r+\delta} n_{x^*R} \left[ n_{Px^*} - \frac{n_{Cx^*}}{n_{CQ}} \right] \\ - \frac{1}{(r+\delta)} n_{x^*R} \frac{n_{Cx^*}}{n_{CQ}} (1+\lambda) \cdot \frac{\Sigma \left[ P_{it} Q_{it} \cdot \frac{\partial f}{\partial Q_t} \cdot \frac{Q_{it}}{P_{it}} \right]}{\Sigma P_{it} Q_{it}}$$

If we let  $A = \frac{\Sigma \left[ P_{it} Q_{it} \cdot \frac{\partial f}{\partial Q_t} \cdot \frac{Q_{it}}{P_{it}} \right]}{\Sigma P_{it} Q_{it}}$

and define  $\bar{P}_t = \frac{\Sigma P_{it} Q_{it}}{Q_t}$ , the industry average price,

$$\text{then } A = \frac{\Sigma \left[ P_{it} Q_{it} \cdot \frac{\partial f}{\partial Q_t} \cdot \frac{Q_{it}}{P_{it}} \right]}{\bar{P}_t Q_t} \\ = \Sigma \frac{\partial f}{\partial Q_t} \cdot \frac{Q_t}{\bar{P}_t} \cdot \frac{Q_{it}^2}{Q_t^2}$$

We can define  $\frac{\partial f}{\partial Q_t} \cdot \frac{Q_t}{\bar{P}_t}$  as the inverse of the market elasticity

of demand,  $\frac{1}{\eta_{Dt}}$ , and  $\Sigma \frac{Q_{it}^2}{Q_t^2} = H_t$ , the Herfindahl index of

concentration. This yields our final equation (vi)

$$\frac{R_t}{\sum P_{it} Q_{it}} = \frac{1}{r+\delta} \cdot \eta_{x^*R} \left[ \eta_{Px^*} - \frac{\eta_{Cx^*}}{\eta_{CQ}} \right] - \frac{1}{(r+\delta)} \eta_{x^*R} \cdot \frac{\eta_{Cx^*}}{\eta_{CQ}} \cdot \frac{(1+\lambda)}{\eta_{Dt}} \cdot H_t \quad (vi)$$

which we can write as

$$\frac{R_t}{P_t Q_t} = a + b H_t \quad (vii)$$

We may note that in the case of pure product innovation,  $\eta_{Cx^*} = 0$  and thus  $b = 0$ , i.e. the R & D/Sales ratio is independent of  $H$ . This gives a good guide to how the model works - the amount of R & D expenditure undertaken depends on its profitability and as R & D affects costs, it is the more profitable the higher is the price cost margin. This in turn depends on  $H$ , and thus R & D is related to the Herfindahl index of concentration.

This analysis indicates that the appropriate measure of concentration to use in the relationship between R & D and market structure is the Herfindahl index. Moreover as we can see from (vi) the coefficients  $a$  and  $b$  are made up of industry specific elasticities and parameters and we do not expect these to be the same in each industry. In the next section we deal with the problems involved in empirically estimating from cross section data a model with coefficients that vary in this way.

### III. Data and Estimation Problems.

There are two major data problems. The first is obtaining R & D data at a low level of aggregation and the second is obtaining estimates of the Herfindahl. The Herfindahl indices used were based on those of Sawyer [15], as amended and updated by Waterson<sup>(1)</sup> and are based on the concentration of employment rather than output. The R & D data available has been taken from HMSO [5] and is, but for a few observations, at SIC Order level. It was only possible to find data on both R & D and the Herfindahl for one year (1968) and thus we have a single cross section of 16 observations from which to estimate. Because of the level of aggregation of this data, we use a second data set utilising information on the employment of scientists and technologists (ST) (this is derived from unpublished material collected for the Department of Employment L7A Survey) at MLH level. Again it is a single cross section for 1968. We call these data sets 1 and 2.

With data set 1, the data on total R & D expenditure (TR) can be supplemented by data on the share of this expenditure provided by government (G). There is no clear argument as to whether the relevant variable should be TR or TR-G. i.e. is government just providing funds for R & D that would be undertaken anyway or is it supplementing the amount of R & D expenditures? Utilising data on sales taken from the Census of Industrial production for 1968, we thus consider regressions on data set 1 of the form,

$$\frac{TR}{PQ} i = \alpha + \beta H_i + U_i \quad \text{and} \quad \frac{TR-G}{PQ} i = \alpha_1 + \beta_1 H_i + U_i$$

---

(1) Whom we wish to thank for providing his data.

and on data set 2 of the form,

$$\frac{ST_i}{pQ_i} = \alpha_2 + \beta_2 H_i + U_i$$

In considering the estimation problem, we rewrite equation (vi) or (vii) in the form

$$Y_i = a_i + b_i X_i, \quad i = 1, \dots, n, \quad (\text{viii})$$

where  $Y = R/pQ$ ,  $X = H$  and the parameters  $a_i$ ,  $b_i$  vary across industries.

Introducing an equation error  $U_i$  and writing  $a_i = a + V_i$ ,  $b_i = b + W_i$ , we can specify a simple linear regression as

$$Y_i = a + b X_i + (U_i + V_i + W_i X_i).$$

Fitting this relationship by least squares will then provide unbiased estimates of  $a$  and  $b$  provided  $U_i$  has zero mean and is uncorrelated with  $X_i$ . The parameters  $a$  and  $b$  in this model have an interpretation as weighted averages of the individual industry coefficients  $a_i$ ,  $b_i$ ,  $i = 1, \dots, n$ .

However, following Klein [8] unbiasedness requires the conditions

$$\sum V_i + \sum W_i X_i = 0$$

$$\sum V_i X_i + \sum W_i X_i^2 = 0.$$

Solving these equations for  $a$  and  $b$  we find

$$a = \frac{\sum X_i^2 (\sum a_i + \sum b_i X_i) - \sum X_i (\sum a_i X_i + \sum b_i X_i^2)}{n \sum X_i^2 - (\sum X_i)^2}$$

$$b = \frac{n(\sum a_i X_i + \sum b_i X_i^2) - \sum X_i (\sum a_i + \sum b_i X_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

It is not possible, therefore, given this specification to interpret the intercept and slope as weighted averages, respectively, of the individual intercepts and slopes. In particular the slope coefficient,  $b$ , depends not only on all the individual slopes,  $b_1, \dots, b_n$ , but also on the intercepts,  $a_1, \dots, a_n$ . Thus if we specify the model in terms of parameters which vary systematically from industry to industry, we are faced with the problem, without further assumptions, that the estimates do not bear any simple relation to them.

One way out of this difficulty which may be justified in some circumstances is to treat coefficients, not as constants, but as realisations of random variables whose distributions depend on a few parameters. The estimation problem in such random coefficient regression models then becomes that of finding suitable values for these parameters, in particular the means.

Writing  $E(a_i) = \alpha$ ,  $E(b_i) = \beta$ , equation (viii) becomes

$$Y_i = \alpha + \beta X_i + U_i$$

where the error term  $U_i = (a_i - \alpha) + (b_i - \beta) X_i$ . If we let

$E(a_i - \alpha)^2 = \sigma_a^2$ ,  $E(b_i - \beta)^2 = \sigma_b^2$ ,  $E(a_i - \alpha)(b_i - \beta) = \sigma_{ab}$ , then the variance of the error term is

$$E(U_i^2) = \sigma_a^2 + 2\sigma_{ab}X_i + \sigma_b^2X_i^2.$$

Provided the industry coefficients  $a_i$ ,  $b_i$  are uncorrelated with the independent variable, market concentration, then least squares estimates are unbiased. Ignoring the heteroscedasticity, however, means that they are inefficient in general and we would expect their estimated standard errors to be biased. The standard error of the estimated slope coefficient will be negatively biased since the error variance is positively related to the independent variable. Goldfeld and Quandt [2] present some illustrative calculations of the inefficiency and bias of the least squares estimates on different assumptions about the strength of heteroscedasticity and the behaviour of the independent variable. The usefulness of least squares was shown to depend very much on the skewness of the independent variable; the more skewed to the right, the less the efficiency of least squares and the greater the bias in the standard errors. In our case the independent variable is considerably skewed and therefore there might be a non negligible payoff to making explicit allowance for the heteroscedasticity. Table I contains some calculations of the bias and inefficiency for one of our independent variables, under different assumptions about heteroscedasticity. The model is  $Y_i = \alpha + \beta X_i + U_i$  with  $\text{Var}(U_i) = a + bX_i + cX_i^2$ .

Table I

a	b	c	Efficiency <sup>(1)</sup>		Bias in estimated standard error <sup>(2)</sup>	
			$\alpha$	$\beta$	$\alpha$	$\beta$
100	0	0	1	1	0	0
100	1000	0	0.99	0.98	0.07	-0.15
100	5000	0	0.92	0.93	0.22	-0.29
100	10000	0	0.85	0.91	0.29	-0.33
100	0	1000	1	1	0.01	-0.03
100	0	10000	0.99	0.97	0.06	-0.19
100	0	100000	0.84	0.82	0.33	-0.42
100	0	500000	0.57	0.71	0.54	-0.48
0	0	500000	0.15	0.58	0.65	-0.49
0	0	10000	0.15	0.58	0.66	-0.49
100	1000	10000	0.98	0.95	0.12	-0.24
100	10000	100000	0.76	0.84	0.39	-0.41
100	15000	500000	0.57	0.75	0.52	-0.46

(1) The ratio of standard errors:

$$\frac{s.e.(\tilde{\alpha})}{s.e.(\hat{\alpha})}, \quad \frac{s.e.(\tilde{\beta})}{s.e.(\hat{\beta})} \quad \text{where } \hat{\alpha}, \hat{\beta} \text{ are OLS estimates and } \tilde{\alpha}, \tilde{\beta} \text{ GLS}$$

(2) The bias is expressed as a proportion of the true standard error:

$$\frac{\text{est.s.e.}(\hat{\alpha}) - s.e.(\hat{\alpha})}{s.e.(\hat{\alpha})}, \quad \frac{\text{est.s.e.}(\hat{\beta}) - s.e.(\hat{\beta})}{s.e.(\hat{\beta})}$$

It can be seen from the table that there may be a considerable gain from using an estimator which specifically recognises the heteroscedasticity present in the random coefficient model. On the other hand, of course, this would mean estimating a larger number of parameters, in this case five as against two, and with only a small sample it may prove difficult to achieve a sufficient degree of precision.

#### IV. Estimation and testing of a random coefficient model

Random coefficient regression models have been considered by a number of authors and a survey is provided by Swamy /16/. Two main approaches have been adopted: a two stage Aitken estimator with heteroscedasticity parameters derived from least squares residuals and maximum likelihood in which all parameters are estimated simultaneously. A problem common to both is the necessity to ensure that the heteroscedasticity pattern which emerges in the course of estimation is feasible in terms of the random coefficient hypothesis.

Writing the model as in equation (viii)<sup>(1)</sup>

$$Y_i = a_i + b_i X_i$$

with  $E(a_i, b_i) = (\alpha, \beta)$  and

$$E \begin{bmatrix} a_i - \alpha \\ b_i - \beta \end{bmatrix} (a_i - \alpha \quad b_i - \beta) = \Sigma = \begin{bmatrix} \sigma_1 & \sigma_3 \\ \sigma_3 & \sigma_2 \end{bmatrix}$$

---

(1) Since the equation error is indistinguishable from the random intercept we omit it and interpret  $\sigma_1$  as the sum of the variances of  $a_i$  and the error term.

is equivalent to a heteroscedastic regression with constant coefficients written as

$$Y_i = \alpha + \beta X_i + U_i$$

where  $E(U_i) = 0$  and  $E(U_i^2) = \sigma_1 + 2\sigma_3 X_i + \sigma_2 X_i^2$ .

A two stage Aitken estimator can be found using an estimate,  $\hat{\Sigma}$ , based on residuals from a first stage least squares regression, to compute a weighted regression

$$W_i Y_i = \alpha W_i + \beta W_i X_i + W_i U_i$$

where  $W_i = (\hat{\sigma}_1 + 2\hat{\sigma}_3 X_i + \hat{\sigma}_2 X_i^2)^{-\frac{1}{2}}$ .

The specification of the model requires  $\hat{\Sigma}$  to be positive definite or positive semi definite. Hildreth and Houck [4] considered the case in which  $\Sigma$  is diagonal where the only problem is to ensure the nonnegativity of the estimated variances. They use two methods of incorporating this a priori information, a truncated estimator where negative estimated variances are given the value of zero, and an estimator based on quadratic programming. Both of these are somewhat unwieldy and are unavailable in our case since we must allow for a non zero covariance. This implies a nonlinear restriction of the form  $\sigma_3^2 \leq \sigma_1 \sigma_2$ .

The same problem is present in the maximum likelihood approach. On the assumption that  $(a_i, b_i)$  is bivariate normal, the log likelihood function is written

$$L = -\frac{1}{2} \sum_{i=1}^n \left[ \log (\sigma_1 + 2\sigma_3 X_i + \sigma_2 X_i^2) + \frac{U_i^2}{\sigma_1 + 2\sigma_3 X_i + \sigma_2 X_i^2} \right] \quad (IX)$$

Unrestricted maximum likelihood estimates can be found by maximising (IX) simultaneously with respect to all five parameters. Restrictions on  $\Sigma$  can be imposed explicitly using Lagrange multipliers. Rubin /14/ considered the case of a diagonal  $\Sigma$  and imposed nonnegativity restrictions in this way but found the resulting likelihood equations difficult to solve. Again this approach seems somewhat cumbersome, particular if we wish to allow for a nonzero covariance.

The problem becomes much more tractable, however, if we can impose the restrictions on  $\Sigma$  directly. If we reparameterise  $\Sigma$  in such a way that it is always positive definite or semidefinite, we can maximise the likelihood function directly using standard numerical optimisation techniques. A simple way of doing this is to write  $\Sigma = \Omega\Omega'$  where  $\Omega$  is lower triangular  $\Omega = \begin{bmatrix} \omega_1 & 0 \\ \omega_2 & \omega_3 \end{bmatrix}$ . The elements of  $\Sigma$  are given in terms of those of  $\Omega$  by  $\sigma_1 = \omega_1^2$ ,  $\sigma_2 = \omega_2^2 + \omega_3^2$ ,  $\sigma_3 = \omega_1\omega_2$  and if all the elements of  $\Omega$  are non zero then  $\Sigma$  is positive definite. The three cases in which  $\Sigma$  is semidefinite correspond to certain elements of  $\Omega$  being equal to zero. The cases  $\sigma_1 = 0$  and  $\sigma_2 = 0$  correspond to  $\omega_1 = 0$  and  $\omega_2 = \omega_3 = 0$  respectively and the case  $\sigma_1 \sigma_2 = \sigma_3^2$  to  $\omega_3 = 0$ . (1)

---

(1) This approach has some points of similarity with that of Nelder [12].

Having obtained point estimates of the parameters, hypothesis tests and confidence intervals can be constructed either by evaluating the information matrix and thus finding approximate standard errors, or by the likelihood ratio approach. In this study we are concerned with tests of the mean regression coefficients and also a test of the random coefficient hypothesis as such. The former are carried out using standard errors estimated from weighted regression using the maximum likelihood values for  $\Sigma$  to generate weights. This amounts to the same as using the information matrix to obtain standard errors for the estimates of  $\alpha$  and  $\beta$ .

A test of the random coefficient hypothesis is based on a comparison of the likelihood function at its global maximum with its value when the regression parameters are estimated under the assumption that they are constant. The latter case corresponds to the standard linear regression model with an additive (homoscedastic) error term and implies restrictions on the elements of  $\Sigma$ . Since the behaviour of the error term is indistinguishable from random variation in the intercept we must allow  $\sigma_1$  to be nonzero but impose the restrictions  $\sigma_2 = \sigma_3 = 0$ . In terms of the elements of  $\Omega$  this amounts to  $\omega_2 = \omega_3 = 0$ , two restrictions on the parameter space. The (approximate) likelihood ratio test is constructed by evaluating (ix) in both cases.

$$\text{If } \tilde{L} = -\frac{1}{2} \sum_{i=1}^n \left\{ \log \left[ \tilde{\omega}_1^2 + 2 \tilde{\omega}_1 \tilde{\omega}_2 X_i + (\tilde{\omega}_2^2 + \tilde{\omega}_3^2) X_i^2 \right] + \frac{(Y_i - \tilde{\alpha} - \tilde{\beta} X_i)^2}{\tilde{\omega}_1^2 + 2 \tilde{\omega}_1 \tilde{\omega}_2 X_i + (\tilde{\omega}_2^2 + \tilde{\omega}_3^2) X_i^2} \right\} \quad (\text{x})$$

for the unrestricted maximum and

$$\hat{L} = -\frac{n}{2} \log \hat{\sigma}_1 - \frac{n}{2}$$

where 
$$\hat{\sigma}_1 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2,$$

for the value using least squares estimates, then the test statistic is  $2(\tilde{L} - \hat{L})$  which is asymptotically distributed as  $\chi^2$  with 2 degrees of freedom. This test allows us to employ a mixed estimator in which least squares estimates are used whenever  $2(\tilde{L} - \hat{L}) < \chi^2_{\alpha}$  and maximum likelihood estimates taken otherwise, at the 100% level. Monte Carlo experiments which provide evidence supporting the use of this kind of mixed estimator in the general heteroscedasticity case are reported in Goldfeld and Quandt [2]. Both ordinary least squares and maximum likelihood estimates are presented below in Section V.

Besides these we also present two stage estimates obtained without the normality assumption. Denoting the least squares residuals by  $e_i$  the elements of  $\Omega$  (and hence  $\Sigma$ ) were estimated by the regression

$$e_i^2 = \omega_1^2 + 2 \omega_1 \omega_2 X_i + (\omega_2^2 + \omega_3^2) X_i^2 + \text{error},$$

i.e. by numerically minimising

$$\sum_{i=1}^n (e_i^2 - \omega_1^2 - 2 \omega_1 \omega_2 X_i - (\omega_2^2 + \omega_3^2) X_i^2)^2. \quad (\text{xi})$$

Having estimates the  $\omega$ 's in this way values for  $\alpha$  and  $\beta$  were found by weighted regression after using these estimates to compute weights.

This approach was used, in addition to maximum likelihood, to meet the suggestion that the latter requires very rigid assumptions about the specification of the model. Since the two stage method just outlined

does not require any distributional assumptions it might be expected to be relatively robust in this respect.

#### V. Results

The estimates are presented in Table II. The ML point estimates were obtained using a Powell algorithm applied to the likelihood function  $(x)$  and the  $t$  values and  $R^2$  found conditionally on the ML values for  $\Sigma$  by weighted regression. The first stage of the two stage estimator, to find estimates of the  $\sigma$ 's, was also carried out using a Powell algorithm to minimise  $(xi)$ . For the OLS results two sets of  $t$  values are shown. The first set were computed using biased standard error estimates and the second using consistent estimates of the standard errors based on the ML values for  $\Sigma$ .

The last column of the table contains the values of the likelihood ratio test statistic. The critical values of  $\chi_2^2$  at the 5% level and 1% levels are, respectively, 5.99 and 9.21. We can therefore only reject the constant parameter hypothesis in the case of equation (1) when the LR is 14.17. Our mixed estimation procedure is therefore to take the ML estimates for equation (1) and the OLS estimates for the other three equations. In terms of our theory these four estimates indicate that there is no significant relationship between the R & D/sales ratio and the level of concentration. The only situation where a significant relationship is found (OLS estimate of equation (1)) is shown to be due to an error in estimation and is refuted using the ML estimates. These results, it might have been thought, were due to the level of aggregation of the data with regard to equations (1) and (2), but no improvement is observed in equations (3) and (4). Moreover comparing (2) to (1) and (4) to (3), allowance for government contribution and the removal of

TABLE II

Dependent variable <sup>(1)</sup>	method	$\alpha$	$\beta$	$R^2$	$\sigma_1$	$\sigma_2$	$\sigma_3$	LR
1. TR/SA <sub>68</sub>	OLS	5.44 (0.40) (0.63)	635.6 (2.38) (1.35)	0.29				
	ML	9.49 (1.75)	524.7 (1.38)	0.12	106	483388	7167	14.17
	Two Stage	15.06 (1.51)	-346.9 (-0.99)	0.066	829	1030225-29232		
2. TR-G/SA <sub>68</sub>	OLS	11.7 (2.58) (2.23)	78 (0.88) (1.12)	0.05				
	ML	9.5 (1.91)	48.4 (0.81)	0.04	326	49366	-4012	4.31
	Two Stage	13.71 (2.15)	-291.7 (-1.30)		$9 \times 10^5$	$10 \times 10^9$	$-3 \times 10^8$	
3. ST/SA <sub>68</sub>	OLS	5.61 (2.09) (1.92)	13.5 (0.35) (0.47)	0.009				
	ML	7.1 (2.41)	-12 (-0.46)	0.016	76	1490	-337	1.05
	Two Stage	5.96 (2.16)	6.89 (0.20)	0.002	53	140.2	86.2	
4. ST/SA <sub>68</sub> <sup>(2)</sup>	OLS	2.34 (1.47) (2.19)	33.3 (1.54) (1.06)	0.18				
	ML	2.0 (2.42)	40.7 (1.46)	0.17	1.7	1709	53.6	5.82
	Two Stage	2.06 (1.72)	39.1 (1.71)	0.21	7.2	302	46.6	

(1) SA are fm; TR are £'000s; ST are units

(2) Same as equation 3 but excluding observations on scientific instruments and electronics, on the hypothesis that these are industries that will be using large numbers of scientists and technologists in non R & D activities.

outlying observations lead to no significant improvements in the results. We cannot therefore accept the hypothesis of a positive relationship between the industry R & D/sales ratio and the degree of concentration. In terms of our original model either (a) firms do not have reaction functions as we have assumed or (b) in equation (vi) the elasticities in the coefficient on H seems on average to take a value of zero. Our conclusions thus lend further support to the statement of Kennedy and Thirlwall [7, p.61] that,

"the evidence appears to be heavily weighted against the hypothesis that a necessary condition for technological change and progressiveness is that firms should be large scale and dominate the market in which they operate."

## VI. Conclusions

We have described a maximum likelihood estimator for a simple regression model with random coefficients and applied it to the relationship between R & D expenditure and market structure. The rationale for using a random coefficients model in this case is that the theory predicts a different equation for each industry and empirical estimation using a cross section is only meaningful if we consider estimating the average parameter values. A feature of the estimator used is its explicit allowance for non zero covariances between random parameters, required in this application because of a theoretical functional dependence between the slope and the constant of the relationship in each industry.

The maximum likelihood approach depends on two crucial assumptions. Firstly the normality of the distribution of the random parameters within the sample is difficult to justify a priori. However results obtained without making this assumption, using a two stage Aitken approach (the two stage estimates) are equally negative and on that basis we might regard our inferences as being reasonably robust to non-normality. Secondly we assume that the degree of concentration, as measured by the Herfindahl, is exogenous in the sense of being independent of the stochastic component of the model. Specifically our approach requires the independent variable to be uncorrelated in the limit with both random coefficients. Since the coefficients may be functions of industry parameters the assumptions we have made, noted in the footnote on p.5 that are necessary to achieve independence, are very strong and we may

therefore be criticised on these grounds. With so few observations, however, only the simplest models are feasible and we can regard our single equation approach as a necessary first step while recognising its inadequacies.

The smallness of the sample used here is, of course, such that any gain from using a random coefficients estimator over least squares is likely to be very small and perhaps we ought properly to regard our estimates as only illustrative (specifically, our variance and covariance estimates are highly imprecise). Applied to a large sample, however, covering a larger number of industries, the random coefficients approach might be a suitable estimator for models of this type. (1)

---

(1) Monte Carlo experimental results reported by Goldfeld and Quandt/2/ indicate that the efficiency of ML relative to OLS increases markedly with sample size beyond 30.

## References

1. K.Cowling and M.Watson, "Price-Cost Margins and Market Structure", Economica, forthcoming.
2. S.Goldfeld and R.Quandt, Nonlinear Methods in Econometrics, North-Holland, 1973.
3. D.M.Grether, "Market Structure and R & D", California Institute of Technology Working Paper, 1974.
4. C.Hildreth and J.P.Houck, "Some Estimators for a Linear Model with Random Coefficients", Journal of the American Statistical Association, 63, (1968), pp.584-95.
5. H.M.S.O. Studies in Official Statistics No.4., HMSO, London, 1974.
6. M.I.Kamien and N.L.Schwartz, "Market Structure and Innovation: A Survey", Journal of Economic Literature, March 1975.
7. C.Kennedy and A.P.Thirlwall, "Technical Progress: A Survey", Economic Journal, March 1972.
8. L.R.Klein, A Textbook of Econometrics, 2nd edition, Prentice-Hall, 1974.
9. J.J.Lambin, P.A.Naert, A.Bultez. "Optimal Marketing Behaviour in Oligopoly", European Economic Review, 1975, pp.105-128.
10. S.Nickell and D.Metcalf, Monopolistic Industries and Monopoly Profits, Mimeo.
11. D.Needham, "Market Structure and Firms' R & D Behaviour", Journal of Industrial Economics, June 1975.
12. J.A.Nelder, "Regression, Model-Building and Invariance", Journal of the Royal Statistical Society, Series A, 131, (1968), pp.303-15.
13. M.Nerlove and K.J.Arrow, "Optimal Advertising Policy under Dynamic Conditions", Economica, 1962.
14. H.Rubin, "Note on Random Coefficients", Chapter XIX in T.C.Koopmans (ed.), Statistical Inference in Dynamic Economic Models, New York, John Wiley and Sons, 1950.
15. M.C.Sawyer, "Concentration in British Manufacturing Industry", Oxford Economic Papers, November 1971.
16. P.A.V.B.Swamy, Statistical Inference in Random Coefficient Regression Models, New York, Springer-Verlag, 1971.

### Appendix

Given the Lagrangean, (1), with  $x_{it} = x_{it}(R_{it})$ , and the firms demand function  $P_{it} = f(Q_{1t} + Q_{2t} \dots Q_{nt}) + g(x_{1t}^*, x_{2t}^* \dots x_{nt}^*)$

$$L = \int_0^{\infty} e^{-rt} \left[ P_{it} Q_{it} - C_{it}(Q_{it}, x_{it}^*) - R_{it} \right] + \lambda(t) (x_{it}^* - x_{it} + \delta_1 x_{it}^*) \quad (i)$$

the necessary Euler conditions for a maximum are,

$$\frac{\partial L}{\partial Q_{it}} = e^{-rt} \left[ P_{it} + Q_{it} \frac{\partial P_{it}}{\partial Q_{it}} - \frac{\partial C_{it}}{\partial Q_{it}} \right] = 0, \quad (ii)$$

$$\frac{\partial L}{\partial R_{it}} = -e^{-rt} - \lambda(t) \frac{\partial x_{it}^*}{\partial R_{it}} = 0, \quad (iii)$$

$$\frac{\partial L}{\partial x_{it}^*} - \frac{d(\partial L / \partial x_{it}^*)}{dt} = e^{-rt} \left[ Q_{it} \frac{\partial P_{it}}{\partial x_{it}^*} - \frac{\partial C_{it}}{\partial x_{it}^*} \right] + \delta_1 \lambda(t) - \frac{d\lambda(t)}{dt} = 0. \quad (iv)$$

$$\text{Thus } e^{-rt} \left[ Q_{it} \frac{\partial P_{it}}{\partial x_{it}^*} - \frac{\partial C_{it}}{\partial x_{it}^*} \right] - (\delta_1 + r) \left[ \frac{e^{-rt}}{\frac{\partial x_{it}^*}{\partial R_{it}}} \right] = 0 \quad (v)$$

after substitution from (iii), and assuming

$$d \left( \frac{\partial x_{it}^*}{\partial R_{it}} \right) / dt = 0.$$

From (ii) we derive (vi)<sup>(1)</sup>

$$P_{it} Q_{it} + P_{it} Q_{it} \cdot \frac{\partial P_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{P_{it}} - \frac{\partial C_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{C_{it}} \cdot C_{it} = 0 \quad (\text{vi})$$

From (v) we obtain that

$$R_{it} = \frac{1}{\delta_i + r} \frac{\partial x_{it}}{\partial R_{it}} \cdot \frac{R_{it}}{x_{it}} \cdot \frac{x_{it}}{x_{it}^*} \left[ P_{it} Q_{it} \cdot \frac{\partial P_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{P_{it}} - \frac{\partial C_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{C_{it}} \cdot C_{it} \right] \quad (\text{vii})$$

Given that (2)  $x_{it}^* = x_{it} + \delta x_{it}^*$ ,

$$\frac{\partial x_{it}}{\partial R_{it}} = \frac{\partial x_{it}^*}{\partial R_{it}}$$

Thus

$$R_{it} = \frac{1}{\delta_i + r} \cdot \frac{\partial x_{it}^*}{\partial R_{it}} \cdot \frac{R_{it}}{x_{it}^*} \left[ P_{it} Q_{it} \cdot \frac{\partial P_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{P_{it}} - \frac{\partial C_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{C_{it}} \cdot C_{it} \right] \quad (\text{viii})$$

Substituting for  $C_{it}$  from (vi) into (viii).

- (1) This equation leads to a generalisation of the Cowling and Waterson [1] result for they assume  $\frac{\partial C_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{C_{it}}$  to equal unity.
- (2) This enables us to reduce from long period to short period conditions and follows the procedure of Cowling in "Optimality in firm's advertising policies", in Cowling (ed.), Market Structure and Corporate Behaviour, Gray Mills, 1972, London.

$$R_{it} = \frac{1}{r + \delta_i} \cdot \frac{\partial x_{it}^*}{\partial R_{it}} \cdot \frac{R_{it}}{x_{it}^*} \cdot P_{it} \cdot Q_{it} \left[ \frac{\partial P_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{P_{it}} \right]$$

$$- \frac{\partial C_{it}}{\partial x_{it}^*} \cdot \frac{x_{it}^*}{C_{it}} \left[ 1 + \frac{\partial P_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{P_{it}} \right]$$

$$\frac{\partial C_{it}}{\partial Q_{it}} \cdot \frac{Q_{it}}{C_{it}} \quad (ix)$$

Equation (ix) is that used in the main text.