A Simple Test for Heteroscedasticity and Random Coefficient Variation
Author(s): T. S. Breusch and A. R. Pagan
Reviewed work(s):
Source: *Econometrica,* Vol. 47, No. 5 (Sep., 1979), pp. 1287-1294
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1911963
Accessed: 12/10/2012 08:11

A SIMPLE TEST FOR HETEROSCEDASTICITY AND RANDOM COEFFICIENT VARIATION

BY T. S. BREUSCH AND A. R. PAGAN[1]

A simple test for heteroscedastic disturbances in a linear regression model is developed using the framework of the Lagrangian multiplier test. For a wide range of heteroscedastic and random coefficient specifications, the criterion is given as a readily computed function of the OLS residuals. Some finite sample evidence is presented to supplement the general asymptotic properties of Lagrangian multiplier tests.

1. INTRODUCTION

IN SOME APPLICATIONS of the general linear model, the usual assumptions of homoscedastic disturbances and fixed coefficients may be questioned. When these requirements are not met, the loss in efficiency in using ordinary least squares (OLS) may be substantial and, more importantly, the biases in estimated standard errors may lead to invalid inferences. This has caused a number of writers to propose models which relax these conditions and to devise estimators for their more general specifications, e.g., Goldfeld and Quandt [8] for heteroscedasticity and Hildreth and Houck [11] for random coefficients. However, because the effect of introducing random coefficient variation is to give the dependent variable a different variance at each observation, models with this feature can be considered as particular heteroscedastic formulations for the purpose of detecting departure from the standard linear model.

A test for heteroscedasticity with the same asymptotic properties as the likelihood ratio test in standard situations, but which can be computed by two least squares regressions, thereby avoiding the iterative calculations necessary to obtain maximum likelihood estimates of the parameters in the full model, is considered in this paper. The approach is based on the Lagrangian multiplier (LM) test of Aitchison and Silvey [1, 20] which is also known as Rao's efficient score test [18, p. 417]. This statistic is obtained from the results of maximizing the likelihood subject to the parameter constraints implied by the null hypothesis and can be computed either from the Lagrangian multipliers corresponding to the constraints as in [1] or from the first order conditions as in [18].[2] Asymptotic equivalence of the LM test with the likelihood ratio procedure is shown in some detail by Silvey [20]. The test proposed in this paper is "constructive" in the sense of [9, p. 85] because a specific form of heteroscedasticity is distinguished as the alternative to the null hypothesis of homoscedasticity. However, it will be seen that the same LM statistic is appropriate for a fairly wide class of alternative hypotheses.

There are four sections to the paper. In Section 2 the general framework is set out and the statistic is derived, Section 3 considers finite sample properties, and general comments are made in the concluding Section 4.

---

[1] We would like to thank unknown referees for their comments.

[2] Similar ideas have been used in other areas, e.g., the Durbin [5] *h*-statistic for autocorrelation in models with lagged dependent variables as regressors can be derived as an LM statistic.

2. THE TEST STATISTIC

Consider the linear model

$$(1) \qquad y_t = x_t'\beta + u_t \qquad\qquad\qquad (t = 1, \ldots, N)$$

where β is a $(k \times 1)$ vector of coefficient parameters and the disturbances u_t are normally and independently distributed with mean zero and variance

$$(2) \qquad \sigma_t^2 = h(z_t'\alpha).$$

Here function $h(\cdot)$, which is not indexed by t, is assumed to possess first and second derivatives, α is a $(p \times 1)$ vector of unrestricted parameters functionally unrelated to the β coefficients, and the first element in z_t is unity. This allows the null hypothesis of homoscedasticity to be written as

$$H_0: \alpha_2 = \ldots = \alpha_p = 0$$

for then $z_t'\alpha = \alpha_1$ so that $\sigma_t^2 = h(\alpha_1) = \sigma^2$ is constant. It is also assumed that x_t and z_t are exogenous, obeying the conditions set out in Amemiya [2].

The representation in (2) is sufficiently general to include most of the heteroscedastic models distinguished in the literature. These are usually either

$$\sigma_t^2 = \exp(z_t'\alpha)$$

which has been shown by Harvey [10] to encompass the specifications of [6, 15 and 16], or

$$\sigma_t^2 = (z_t'\alpha)^m$$

with m a prespecified integer as in [7, 8, and 19]. The random coefficient model of Hildreth and Houck [11] and most of its later generalizations, for example the one considered by Swamy and Mehta [21], are of the form $\sigma_t^2 = z_t'\alpha$ where the elements of z_t are obtained from the distinct elements of $x_t x_t'$.[3]

Define the OLS residuals from (1) as $\hat{u}_t$ and the estimated residual variance as $\hat{\sigma}^2 = N^{-1} \Sigma \hat{u}_t^2$. This allows our basic result to be stated in the following theorem.

THEOREM: *For the model* (1) *and* (2) *under the conditions given above, the Lagrangian multiplier statistic for testing* $H_0: \alpha_2 = \ldots = \alpha_p = 0$ *(homoscedastic disturbances) can be found as one half the explained sum of squares in a regression of* $g_t = \hat{\sigma}^{-2} \hat{u}_t^2$ *upon* z_t *and is asymptotically distributed as* χ^2 *with* $(p-1)$ *degrees of freedom when the null hypothesis is true.*[4]

---

[3] Note that there are some heteroscedastic models which do not fit our general formulation, e.g., $\sigma_t^2 = \alpha z_t$ (scalar α and z_t) and $\sigma_t^2 \propto [E(y_t)]^2$. But these specifications do not provide a convenient framework for testing homoscedasticity. In the first example, there is no parametric restriction which gives the null hypothesis as a special case of the general model, and in the second there is no regression without heteroscedasticity because of the implied relationship between α and β.

[4] The statistic is defined as a regression result to give a convenient method of computation which is not meant to imply that the usual criteria for "goodness of fit" of this regression have any meaningful properties.

PROOF: Let $l(\theta)$ be a log likelihood depending on a vector of parameters θ with $d = \partial l/\partial \theta$ as the first derivative (score) vector and $\mathscr{I} = -E(\partial^2 l/\partial \theta \partial \theta')$ as the information matrix. Then following Rao [18, pp. 418–419] the LM statistic for testing the null hypothesis represented by parametric constraints $\phi(\theta) = 0$ is given by

$$LM = \hat{d}' \hat{\mathscr{I}}^{-1} \hat{d}$$

where the hats indicate that the quantities are evaluated with $\hat{\theta}$, the restricted maximum likelihood estimate satisfying $\phi(\hat{\theta}) = 0$. For the case where $\theta' = (\theta'_1 : \theta'_2)$ and the constraints refer to only one of the subsets, say $\phi(\theta_2) = 0$, the vector d may be partitioned conformably as $d' = (d'_1 : d'_2)$ so that $\hat{d}_1 = 0$ from constrained maximum likelihood. If, furthermore, the information matrix is block diagonal between θ_1 and θ_2, $\mathscr{I}_{21} = -E(\partial^2 l/\partial \theta_2 \partial \theta'_1) = 0$, the statistic becomes

$$LM = \hat{d}'_2 \hat{\mathscr{I}}^{-1}_{22} \hat{d}_2$$

where

$$\mathscr{I}_{22} = -E(\partial^2 l/\partial \theta_2 \partial \theta'_2).$$

For the model given by (1) and (2) the log likelihood is

$$l(\beta, \alpha) = -\tfrac{1}{2} N \log(2\pi) - \tfrac{1}{2} \sum_t \log \sigma_t^2 - \tfrac{1}{2} \sum_t \sigma_t^{-2} (y_t - x'_t \beta)^2$$

where $\sigma_t^2 = h(z'_t \alpha)$. The first derivative with respect to the α parameters is

$$d_\alpha = \partial l/\partial \alpha = \tfrac{1}{2} \sum_t h'(s_t) z_t (\sigma_t^{-4} u_t^2 - \sigma_t^{-2})$$

where $s_t = z'_t \alpha$ and $h'(s_t) = \partial h(s_t)/\partial s_t$. It is easily seen that $\mathscr{I}_{\alpha\beta} = -E(\partial^2 l/\partial \alpha \partial \beta') = 0$ so that the LM statistics for testing H_0 will be $\hat{d}'_\alpha \hat{\mathscr{I}}^{-1}_{\alpha\alpha} \hat{d}_\alpha$ where constrained maximum likelihood corresponds to OLS applied to (1). Evaluating the required quantities gives

$$\hat{d}_\alpha = \tfrac{1}{2} [\hat{\sigma}^{-2} h'(\hat{\alpha}_1)] \sum_t z_t (\hat{\sigma}^{-2} \hat{u}_t^2 - 1),$$

$$\hat{\mathscr{I}}_{\alpha\alpha} = \tfrac{1}{2} [\hat{\sigma}^{-2} h'(\hat{\alpha}_1)]^2 \sum_t z_t z'_t,$$

and the statistic is

$$(3) \qquad LM = \tfrac{1}{2} \Big(\sum_t z_t f_t \Big)' \Big(\sum_t z_t z'_t \Big)^{-1} \Big(\sum_t z_t f_t \Big)$$

where $f_t = (\hat{\sigma}^{-2} \hat{u}_t^2 - 1) = g_t - 1$. Alternatively, collect all N observations by defining $Z = (z_1, \ldots, z_N)'$, $f = (f_1, \ldots, f_N)'$, $g = (g_1, \ldots, g_N)'$, and i as an $(N \times 1)$ vector of units. Then $f = (g - i)$, $i'g = N$, $i'f = 0$, and $f'Z(Z'Z)^{-1}Z'i = 0$ because i is the first column of Z. Thus

$$LM = \tfrac{1}{2} f'Z(Z'Z)^{-1}Z'f$$

$$= \tfrac{1}{2} [g'Z(Z'Z)^{-1}Z'g - N^{-1}(i'g)^2]$$

which is one half of the explained sum of squares in the regression of g_t upon z_t (see Goldberger [12, p. 165, eq. (4.21)]).

Let $\hat{u}^2 = \hat{\sigma}^2 g$, let S be the explained sum of squares from the regression of $\hat{u}^2$ against Z, and let $\hat{\delta}$ be the OLS estimates of the coefficients (δ) in this regression. From Amemiya [2, eqs. (5) and (6)] it follows that, *under the null hypothesis*, $N^{\frac{1}{2}}(\hat{\delta} - \delta) \rightarrow N(0, 2\sigma^4(N^{-1}Z'Z)^{-1})$ so that $(2\sigma^4)^{-1}S \rightarrow \chi^2_{p-1}$ as in standard least squares theory. As $LM = (2\hat{\sigma}^4)^{-1}S$ and $\hat{\sigma}^4 \rightarrow \sigma^4$ in probability under H_0, it follows that $LM \rightarrow \chi^2_{p-1}$ in distribution.

3. FINITE SAMPLE PROPERTIES

As with all procedures developed from asymptotic principles, it is desirable to investigate the properties of the statistic based on a finite amount of data. It is difficult to establish the exact small sample distribution by analytical methods as the LM statistic is the ratio of quadratic forms in the OLS residuals $\hat{u}_t^2$, which are dependent gamma variables, and very little work has been done in analyzing such distributions. However, as the distribution of the LM statistic under H_0 can be shown to be independent of any unknown parameters, the Type I error can always be evaluated in the context of a particular model by Monte Carlo methods (to any desired degree of accuracy).

It is useful to explore the finite sample properties in the case $p = 2$, i.e. $\sigma_t^2 = h(\alpha_1 + \alpha_2 z_{2,t})$, in some detail. The LM statistic is then

$$LM = \frac{1}{2}\left[\sum_t (z_{2,t} - \bar{z}_2)^2\right]^{-1}\left[\sum (z_{2,t} - \bar{z}_2)f_t\right]^2$$

$$= \frac{1}{2}\left[\sum_t (z_{2,t} - \bar{z}_2)^2\right]^{-1}\left[\sum (z_{2,t} - \bar{z}_2)g_t\right]^2$$

$$= (\hat{u}'D\hat{u}/\hat{u}'\hat{u})^2$$

where $\bar{z}_2 = N^{-1}\sum_t z_{2,t}$, $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_N)'$, and D is a diagonal matrix with ith diagonal element $\{N(z_{2,i} - \bar{z}_2)/[2\sum_t (z_{2,t} - \bar{z}_2)^2]^{\frac{1}{2}}\}$ for $i = 1, \ldots, N$. Thus for any $c > 0$,

$$\text{pr}\{LM > c\} = \text{pr}\{\hat{u}'D\hat{u}/\hat{u}'\hat{u} > \sqrt{c}\} + \text{pr}\{\hat{u}'D\hat{u}/\hat{u}'\hat{u} < -\sqrt{c}\}.$$

Because each of these terms involves the ratio of quadratic forms in normal variables, Imhof's procedure [13] might be used to compute the *exact* probabilities (see Koerts and Abrahamse [14], for a good discussion of this). The disadvantage with this method is that it does not extend beyond $p = 2$ and an alternative that covers any p is desirable. Such an alternative is available by observing that division of both numerator and denominator by σ^2 does not affect LM but results in $\hat{u}$ being a linear transformation of the standard normal deviates $\sigma^{-1}u_t$. Thus, assuming an investigator has computed a value of $LM = c$ from some model, the exact probability of a Type I error for this value can be estimated by (a) generating n sets of N observations on an n.i.d.(0, 1) variable ϕ_t, (b) forming the statistic as in (3) using $\hat{u} = (I - X(X'X)^{-1}X')\phi$, (c) observing the number of times,

r, that $LM > c$ in these n sets, (d) using r/n as the estimate of the probability of a Type I error. The difference between r/n and the *exact* probability tends to zero as $n \to \infty$, so that this difference can be made arbitrarily small with high probability, and good results seem likely for $n = 5000$. Although this may seem expensive it should be borne in mind that, even when $p = 2$ and Imhof's method is available, exact computation via inversion of the characteristic function involves two numerical integrations, and these also exhibit errors that may be as high as those from simulation if the truncation point is too small or the grid is too large.

Even though there is a way of computing the exact probability of Type I errors for the statistic, it seems likely that most researchers would only use this if the computed LM lay near the critical 5 per cent or 10 per cent significance points of the χ^2_{p-1} and it therefore seems worthwhile assessing the adequacy of such a strategy for $p = 2$ with a particular set of data. Such a choice also enables a comparison of the simulation and Imhof methods, both in terms of accuracy and computational cost. Essentially, there are three important questions to be asked concerning the small sample distribution: How adequate is the χ^2 approximation as an indicator of significance levels? What is the power of the test statistic to reject a false null hypothesis of homoscedasticity? How robust is the test to a misspecified model? In the following experiments attention is centered upon the first two questions.

Rutemiller and Bowers [19] investigated heteroscedasticity in two regression models and it was decided to select the data from these models (but not the same heteroscedastic model) for the experiments. Model I utilizes the "radio set" data while Model II works with the "auto stopping distance" data. Accordingly, there are a maximum of forty-nine observations for Model I and sixty-three for Model II.

Having selected the data it is necessary to decide on a form for the heteroscedasticity. The random coefficient model $y_t = \alpha + \beta_t x_t + u_t$, $\beta_t = \bar{\beta} + \varepsilon_t$, which implies $\sigma_t^2 = \sigma_u^2 + x_t^2 \sigma_\varepsilon^2$, was selected so that $z_{2,t}$ in our experiments will always be x_t^2. The null hypothesis that $\sigma_\varepsilon^2 = 0$ is an interesting one as the value lies on the boundary of the parameter space and, as Chernoff [4] pointed out, the likelihood ratio test would not be χ^2 in such a situation. However, as Chant [3] observes, in this non-standard situation the LM statistic will be χ^2 in large samples.

Table I records the predicted probability of Type I error from the asymptotic theory (column 1) and the exact probabilities for various sample sizes for the two models.

From Table I the adequacy of the asymptotic theory to indicate correct significance levels is rather suspect. Certainly, it would appear that investigators might need to use fairly conservative significance levels. Because of this divergence of asymptotic predictions and small sample results, we examined the ability of the simulation method to provide the user with good approximations to the true probability of Type I error. Table II contains a comparison of the exact probabilities generated by the Imhof method with those from the simulation method (TIME is the C.P.U. time in seconds for computing the whole column on a UNIVAC 1142).

TABLE I

PROBABILITY OF TYPE I ERRORS FOR VARIOUS
SAMPLE SIZES

		Model I			Model II	
$N = \infty$	$N = 20$	$N = 40$	$N = 49$	$N = 20$	$N = 40$	$N = 60$
.7	.691	.698	.710	.701	.690	.695
.5	.449	.460	.487	.497	.484	.490
.4	.312	.324	.362	.393	.381	.388
.3	.172	.183	.229	.287	.280	.286
.2	.058	.066	.104	.180	.180	.185
.1	.018	.022	.033	.078	.084	.088
.05	.010	.013	.021	.034	.040	.042
.02	.005	.008	.012	.013	.016	.017
.01	.003	.005	.009	.007	.008	.009
.005	.002	.004	.006	.004	.005	.005

TABLE II

SIMULATION AND IMHOF METHODS FOR
EVALUATING PROBABILITIES, MODEL I

	$N = 20$		$N = 40$	
	SIM	IMHOF	SIM	IMHOF
	.6970	.6906	.7204	.7096
	.4594	.4488	.4834	.4873
	.3212	.3121	.3608	.3620
	.1756	.1720	.2230	.2293
	.0536	.0579	.1034	.1041
	.0192	.0179	.0352	.0330
	.0106	.0101	.0222	.0207
	.0048	.0050	.0126	.0123
	.0024	.0031	.0088	.0086
	.0016	.0019	.0064	.0062
Time	34.5	71.4	92.0	703.3

Assuming that the Imhof method yields the exact probabilities it is seen that the simulation method provides a reliable guide to the evaluation of these in any applied situation—certainly errors are minor compared to those from the use of the asymptotic theory. It is of interest to note that the simulation method, based on 5000 replications, was considerably faster than the Imhof method.

To assess the power of the statistic it is necessary to specify particular numerical values for the ratio α_2/α_1 (i.e., $\sigma_\varepsilon^2/\sigma_u^2$). Because the model being investigated is a random coefficient one, it seemed sensible to relate power to the coefficient of variation (CV) of β_i, i.e., $\sigma_\varepsilon/\bar{\beta}$. By choosing $\sigma_u^2 = 0.3$ and $\bar{\beta}$ as 0.66 for Model II (roughly the OLS estimates), σ_ε^2 was found for values of CV of 0.1, 1.0, and 10.0, these constituting a reasonable range of randomness in β_i. Table III records the rejection probabilities of the test statistic for the 10 critical values of Table I, as the CV and sample size change.[5]

[5] The tabulated figures differ from power because of the variations in Type I error levels given in Table I.

From Table III it appears that, at least for the range of the CV considered, sample size is the main determinant of power and this is acceptable for sample sizes of forty and above, for those significance levels most commonly in use, i.e., 5 per cent and 10 per cent.

TABLE III

POWER CALCULATIONS FOR MODEL II

| | $N = 20$ | | | $N = 40$ | | | $N = 60$ | |
CV = 0.1	CV = 1.0	CV = 10.0	CV = 0.1	CV = 1.0	CV = 10.0	CV = 0.1	CV = 1.0	CV = 10.0
.935	.943	0.944	0.995	0.996	0.996	1.000	1.000	1.000
.879	.893	0.895	0.988	0.990	0.990	0.999	1.000	1.000
.842	.859	0.861	0.982	0.984	0.985	0.999	0.999	0.999
.793	.815	0.817	0.971	0.974	0.975	0.998	0.998	0.999
.725	.750	0.752	0.950	0.955	0.956	0.996	0.997	0.997
.608	.637	0.640	0.900	0.909	0.910	0.989	0.991	0.991
.498	.528	0.531	0.835	0.848	0.849	0.977	0.980	0.981
.369	.398	0.401	0.734	0.751	0.753	0.952	0.958	0.958
.288	.313	0.316	0.653	0.673	0.672	0.925	0.933	0.934
.220	.242	0.245	0.573	0.592	0.594	0.891	0.902	0.903

4. COMMENTS AND CONCLUSION

1. Although the statistic can be employed only when an alternative is specified, its derivation suggests that the quantity $g_t = \hat{u}_t^2/\hat{\sigma}^2$ is of some importance in tests of heteroscedasticity. Thus, *if one is going to plot any quantity*—a strategy sometimes recommended—it would seem to be more reasonable to plot g_t than quantities such as $\hat{u}_t$.

2. The test statistic is easily extended to systems of equations and in some circumstances can be expressed in a simpler form, e.g., if it is assumed that there is a discrete change in σ_t^2 after n observations, as is common in many cross-sectional studies, then $z_{2,t}$ would be unity from 1 to n and zero elsewhere, from which the statistic becomes

$$\text{LM} = \frac{1}{2}\left[\frac{N}{n(N-n)}\right]\left[\sum_t^n \left(\frac{\hat{u}_t^2}{\hat{\sigma}^2}\right) - n\right]^2.$$

This essentially involves a comparison of the residual sum of squares over the first n periods and the total sample. Unfortunately, even in this special case the distribution of the statistic does not appear tractable analytically.

3. Finally, there is the question of the power of the test statistic versus (say) the likelihood ratio (LR) statistic, i.e., even if the LM statistic did not have good power in small samples, it may be no worse than others and its computational ease might make it more attractive (in the same way that the Durbin h statistic is more attractive than the LR test). Although Peers [17] has demonstrated that the LM

test was biased to $O(N^{-\frac{1}{2}})$, it was impossible to predict which would be the more powerful of the test statistics, and this question probably needs to be taken up in further research.

University of Southampton
and
Australian National University

REFERENCES

[1] AITCHISON, J., AND S. D. SILVEY: "Maximum-Likelihood Estimation and Associated Tests of Significance," *Journal of the Royal Statistical Society*, Series B, 22 (1960), 154–171.
[2] AMEMIYA, T.: "A Note on a Heteroscedastic Model," *Journal of Econometrics*, 6 (1977), 365–370.
[3] CHANT, D.: "On Asymptotic Tests of Composite Hypotheses in Nonstandard Conditions," *Biometrika*, 61 (1974), 291–298.
[4] CHERNOFF, H.: "On the Distribution of the Likelihood Ratio," *Annals of Mathematical Statistics*, 25 (1954), 573–578.
[5] DURBIN, J.: "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica*, 38 (1970), 410–421.
[6] GEARY, R. C.: "A Note on Residual Heterovariance and Estimation Efficiency in Regression," *American Statistician*, 20 (1966), 30–31.
[7] GLEJSER, H.: "A New Test for Heteroskedasticity," *Journal of the American Statistical Association*, 64 (1969), 316–323.
[8] GOLDFELD, S. M., AND R. E. QUANDT: "Some Tests for Homoscedasticity," *Journal of the American Statistical Association*, 60 (1965), 539–547.
[9] ———: *Nonlinear Methods in Econometrics.* Amsterdam: North-Holland, 1972.
[10] HARVEY, A. C.: "Estimating Regression Models with Multiplicative Heteroscedasticity," *Econometrica*, 44 (1976), 461–466.
[11] HILDRETH, C., AND J. P. HOUCK: "Some Estimators for a Linear Model with Random Coefficients," *Journal of the American Statistical Association*, 63 (1968), 584–595.
[12] GOLDBERGER, A. S.: *Econometric Theory.* New York: John Wiley and Sons, 1964.
[13] IMHOF, J. P.: "Computing the Distribution of Quadratic Forms in Normal Variables," *Biometrika*, 48 (1961), 419–426.
[14] KOERTS, J., AND A. P. J. ABRAHAMSE: *On the Theory and Application of the General Linear Model.* Rotterdam: Universitaire Pers Rotterdam, 1969.
[15] LANCASTER, T.: "Grouping Estimators on Heteroscedastic Data," *Journal of the American Statistical Association*, 63 (1968), 182–191.
[16] PARK, R. E.: "Estimation with Heteroscedastic Error Terms," *Econometrica*, 34 (1966), 888.
[17] PEERS, H. W.: "Likelihood Ratio and Associated Test Criteria," *Biometrika*, 58 (1971), 577–587.
[18] RAO, C. R.: *Linear Statistical Inference and Its Applications*, 2nd Ed. New York: John Wiley and Sons, 1973.
[19] RUTEMILLER, H. C., AND D. A. BOWERS: "Estimation in a Heteroscedastic Regression Model," *Journal of the American Statistical Association*, 63 (1968), 552–557.
[20] SILVEY, S. D.: "The Lagrangian Multiplier Test," *Annals of Mathematical Statistics*, 30 (1959), 389–407.
[21] SWAMY, P. A. V. B., AND J. S. MEHTA: "Bayesian and Non-Bayesian Analysis of Switching Regressions and of Random Coefficient Regression Models," *Journal of the American Statistical Association*, 70 (1975), 593–602.