# Testing for Aggregation Bias in Linear Models*

Kevin C. Lee
*Department of Applied Economics, Cambridge University, Cambridge CB3 9DE, UK*

M. Hashem Pesaran
*Trinity College, Cambridge CB2 1TQ, UK*

Richard G. Pierse
*Department of Applied Economics, Cambridge University, Cambridge CB3 9DE, UK*

1990

The problem of aggregation over micro units has had a long tradition in the econometrics literature, stretching back to the pioneering work of Theil (1954). In this literature two issues in particular have attracted attention. The first concentrates on the prediction problem of choosing whether to use macro or micro equations to predict aggregate variables. This issue was raised by Grunfeld and Griliches (1960) and is further addressed in a recent paper by Pesaran et al. (1989) (PPK). In PPK a generalised prediction criterion and a formal statistical test of the hypothesis of perfect aggregation are developed. The present paper considers the second strand in this literature which is concerned with the problem of 'aggregation bias' defined by the deviation of the macro parameters from the average of the corresponding micro parameters. (See for example Theil (1954), Boot and de Wit (1960), Orcutt et al. (1968), Gupta (1971) and Sasaki (1978).) In this paper we develop direct tests of aggregation bias in contrast to the indirect test proposed by Zellner (1962) which tests the hypothesis that all the disaggregated coefficients are equal. We also derive generalised versions of the tests for the case where the parameters of interest are subsets or (possibly non- linear) functions of the full parameter vector. This is particularly relevant when the focus of the analysis is on the long run properties of the aggregate and disaggregate models. Since the tests of the aggregation bias, whether of the type discussed here or the one proposed in Zellner (I962), assume the disaggregate model is correctly specified, in this paper we also develop a Durbin-Hausman type misspecification test of the disaggregate model. Section I sets out the statistical framework and assumptions. Section II develops the aggregation bias tests. Section III derives the Durbin-Hausman type misspecification test of the disaggregate model. Section IV applies these tests to a disaggregate model of employment demand for the United Kingdom taken from PPK.

## I Framework and Assumptions

In order to develop the tests we consider the following general disaggregate model:

$$H_d: \quad \mathbf{y}_t = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{u}_i, \quad i = 1, 2, \ldots, m \tag{1}$$

where $\mathbf{y}_i$ is the $n \times 1$ vector of observations on the dependent variable for the $i$th unit, $\mathbf{X}_i$ is the $n \times k$ matrix of observations on the regressors in (1) for the $i$th unit, $\boldsymbol{\beta}_i$ is the $k \times 1$ vector of the coefficients associated with columns of $\mathbf{X}_i$, and $\mathbf{u}_i$ is the $n \times 1$ vector of disturbances for the $i$th unit. The corresponding aggregate equation that satisfies the Klein-Nataf consistency requirement is given by [1]

$$H_a: \quad \mathbf{y}_a = \mathbf{X}_a\mathbf{b}_a + \mathbf{v}, \tag{2}$$

*Published in *Economic Journal* (1990), Vol. 100 (Conference 1990), pp. 137–150.

[1] See Lovell (1973), and the discussion in PPK (p. 25).

where

$$\mathbf{y}_a = \sum_{i=1}^{m} \mathbf{y}_i, \quad \mathbf{X}_a = \sum_{i=1}^{m} \mathbf{X}_i,$$

and $\mathbf{b}_a$ is the $k \times 1$ vector of macro parameters. The $n \times 1$ disturbance vector $\mathbf{v}$, will be equal to $\mathbf{u}_a = \sum_{i=1}^{m} \mathbf{u}_i$, only if the 'perfect aggregation' condition

$$H_\xi : \boldsymbol{\xi} = \sum_{i=1}^{m} \mathbf{X}_i \boldsymbol{\beta}_i - \mathbf{X}_a \mathbf{b}_a = \mathbf{0}, \tag{3}$$

discussed in detail in PPK, is satisfied. Here we focus on the problem of aggregation bias and develop alternative methods of analysing and formally testing the extent of this bias in economic applications. In what follows we adopt the following assumptions:

**Assumption 1.** *The $n$ elements of the disturbance vector $\mathbf{u}_i = \{u_{it}\}$, have zero means, constant variances and are serially independently distributed. They also satisfy the moment condition*

$$\mathrm{E}\,|u_{it}|^{2+\delta} < \Delta < \infty, \text{ for some } \delta > 0, \text{ and all } t.$$

**Assumption 2.** *The disturbance vectors $\mathbf{u}_i$ are distributed independently of $\mathbf{X}_i$, and $\mathrm{E}(\mathbf{u}_i \mathbf{u}_j{}') = \sigma_{ij}\mathbf{I}_n$, for all $i$ and $j$, $(\sigma_{ii} > 0)$.*

**Assumption 3.** *The matrices $\mathbf{X}_i$ have full rank, the probability limits*

$$\operatorname*{plim}_{n\to\infty}(n^{-1}\mathbf{X}_i{}'\mathbf{X}_i) = \boldsymbol{\Sigma}_{ij}, \quad i, j = a, 1, 2, \ldots, m,$$

*exist, and the $k \times k$ matrices $\boldsymbol{\Sigma}_{ii}$, $i = a, 1, 2, \ldots, m$ are non-singular.*

We also base our tests on the OLS estimates

$$\widehat{\mathbf{b}} = (\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}'\mathbf{y}_a, \quad \widehat{\boldsymbol{\beta}}_i = (\mathbf{X}_i{}'\mathbf{X}_i)^{-1}\mathbf{X}_i{}'\mathbf{y}_i, \quad i = 1, 2, \ldots, m,$$

although, in principle, the tests proposed below can also be constructed using the more efficient SURE (Seemingly Unrelated Regression Equations) estimators of $\boldsymbol{\beta}_i$, due to Zellner (1962).

# II Direct Tests of Aggregation Bias

The problem of 'aggregation bias', as originally discussed by Theil (1954) is defined in terms of the deviations of macro parameters from the averages of the corresponding micro parameters.[2] In the context of the linear disaggregate and aggregate models (1) and (2), the vector of aggregation bias is defined by

$$\boldsymbol{\eta}_\beta = \mathbf{b} - \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{\beta}_i, \tag{4}$$

A test of aggregation bias then involves testing the hypothesis $H_0 : \boldsymbol{\eta}_\beta = \mathbf{0}$. In testing this hypothesis the case where $\mathbf{b}$ is given *a priori* (for example by a 'consensus' view) should be distinguished from the case where $\mathbf{b}$ is defined as the pseudo true value of $\widehat{\mathbf{b}}$ assuming that the disaggregate model is correctly specified. In the former case the relevant statistic for testing the hypothesis $H_0 : \boldsymbol{\eta}_\beta = \mathbf{0}$ is given by

$$q_1 = \left(\mathbf{b} - \frac{1}{m}\sum_{i=1}^{m}\widehat{\boldsymbol{\beta}}_i\right)' \widehat{\boldsymbol{\Omega}}_n^{-1} \left(\mathbf{b} - \frac{1}{m}\sum_{i=1}^{m}\widehat{\boldsymbol{\beta}}_i\right), \tag{5}$$

where $\widehat{\boldsymbol{\Omega}}_n$ represents a consistent estimator of $\boldsymbol{\Omega} = m^{-2}\sum_{i,j=1}^{m} \mathrm{Cov}(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j)$.[3] Under assumptions 1–3 it is easily seen that $q_1$ is asymptotically distributed as $\chi^2_k$ The statistic $q_1$ takes $\mathbf{b}$ as a fixed vector, and tests for the deviation of the average of micro parameters from this fixed vector on the assumption that $H_d$ holds. In practice, however, it is rare that a 'consensus' value for $\mathbf{b}$ or some of its elements is

[2] For empirical analysis of aggregation bias see, for example, the papers by Boot and de Wit (1960), Gupta (1971) and Sasaki (1978).

[3] Note that $\mathrm{Cov}(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j) = \sigma_{ij}(\mathbf{X}_i{}'\mathbf{X}_i)^{-1}(\mathbf{X}_i{}'\mathbf{X}_j)(\mathbf{X}_j{}'\mathbf{X}_j)^{-1}$.

available, and $\mathbf{b}$ needs to be chosen in light of the knowledge of the disaggregate model. When $H_d$ holds the pseudo true value of $\mathbf{b}$ is given by

$$\mathbf{b} = \underset{n\to\infty}{\text{plim}}(\widehat{\mathbf{b}}|H_d) = \sum_{i=1}^{m} \mathbf{C}_i \boldsymbol{\beta}_i, \tag{6}$$

where

$$\mathbf{C}_i = \boldsymbol{\Sigma}_{aa}^{-1}\boldsymbol{\Sigma}_{ai}, \quad i = 1, 2, \ldots, m, \tag{7}$$

satisfy the condition $\sum_{i=1}^{m} \mathbf{C}_i = \mathbf{I}_k$. $\mathbf{I}_k$ is an identity matrix of order $k$.) The matrices $\mathbf{C}_i$ are the probability limits of the coefficients in the OLS regressions of the columns of $\mathbf{X}_i$ on $\mathbf{X}_a$; the 'auxiliary' equations in Theil's terminology. Notice that result (6) holds only when $H_d$ is correctly specified. We will use this result later as the basis of a Durbin-Hausman type test of misspecification of the disaggregate model. For the time being, however, we assume that the disaggregate model $H_d$ is correctly specified and write $H_0$ as

$$H_0; \quad \sum_{i=1}^{m}\left(\mathbf{C}_i - \frac{1}{m}\mathbf{I}_k\right)\boldsymbol{\beta}_i = \mathbf{0}. \tag{8}$$

An indirect, albeit familiar method of testing (8), originally proposed by Zellner (1962), is to test the micro-homogeneity hypothesis

$$H_\beta : \quad \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_m.$$

Testing $H_\beta$ as a method of testing $H_0$ is however rather too restrictive. Although $H_\beta$ implies $H_0$, the reverse is not true. It is possible for $\boldsymbol{\eta}_\beta = \mathbf{0}$ to hold even when the micro-homogeneity hypothesis is rejected. Here we propose a direct test of $H_0$ based on the OLS estimate of $\boldsymbol{\eta}_\beta$, namely

$$\widehat{\boldsymbol{\eta}}_\beta = \widehat{\mathbf{b}} - \frac{1}{m}\sum_{i=1}^{m}\widehat{\boldsymbol{\beta}}_i. \tag{9}$$

Under $H_0$, $\widehat{\boldsymbol{\eta}}_\beta$ is given by

$$\widehat{\boldsymbol{\eta}}_\beta = \sum_{i=1}^{m} \mathbf{P}_i \mathbf{u}_i, \tag{10}$$

where

$$\mathbf{P}_i = (\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}' - \frac{1}{m} - (\mathbf{X}_i{}'\mathbf{X}_i)^{-1}\mathbf{X}_i{}'. \tag{11}$$

This suggests basing a test of $H_0$ on the statistic

$$q_2 = n^{-1}\widehat{\boldsymbol{\eta}}_\beta{}'\widehat{\boldsymbol{\Phi}}_n^{-1}\widehat{\boldsymbol{\eta}}_\beta, \tag{12}$$

$$\widehat{\boldsymbol{\Phi}}_n = n^{-1}\sum_{i,j=1}^{m}\widehat{\sigma}_{ij}\mathbf{P}_i\mathbf{P}_j{}', \tag{13}$$

and $\widehat{\sigma}_{ij}$ is a consistent estimator of $\sigma_{ij}$.[4] Notice that except for the extreme case where $\mathbf{X}_i = m^{-1}\mathbf{X}_a$, matrix $\widehat{\boldsymbol{\Phi}}_n$ will in general be non-singular.

**Theorem 1.** *Suppose*

*(i) The disaggregate model $H_d$ is correctly specified;*

*(ii) Assumptions 1–3 hold;*

*(iii) The matrix $\widehat{\boldsymbol{\Phi}}_n$ defined by (13) and the matrix $n^{-1}(\mathbf{P}_i\mathbf{P}_i{}')$ both are non-singular and also converge in probability to non-singular matrices.*

*Then on the hypothesis of no aggregation bias, $H_0$, the statistic $q_2$ defined in (12) is asymptotically distributed as a chi-squared variate with $k$ degrees of freedom.*

*Proof.* *See the Mathematical Appendix.*

[4] In small samples we suggest using the unbiased (and consistent) estimator of $\sigma_{ij}$ proposed in PPK. (See equation (5.9) in PPK.)

This theorem provides an asymptotic justification for the use of $q_2$ in testing the null hypothesis of no aggregation bias, and holds for $\sigma_{ij} \neq 0$ and $m \geq 2$, but requires $n$, the sample size, to be sufficiently large. This contrasts the asymptotic framework underlying the perfect aggregation test proposed in PPK where $n$ is fixed but m is allowed to increase without bounds.

The test statistics $q_1$ and $q_2$ are applicable when the focus of the analysis is on all the elements of $\boldsymbol{\beta}_i$. In practice, it is often the case that the parameters of interest are subsets or, more generally, (non-linear) functions of $\boldsymbol{\beta}_i$. To deal with such cases we now consider a generalisation of (4) and write the null hypothesis of no aggregation bias as

$$\boldsymbol{\eta}_g = \mathbf{g}(\mathbf{b}) - \frac{1}{m}\sum_{i=1}^{m}\mathbf{g}(\boldsymbol{\beta}_i), \tag{14}$$

where $\mathbf{g}(\boldsymbol{\beta}_i)$ is an $s \times 1$ $(s \leq k)$ vector of known functions of $\boldsymbol{\beta}_i$.

Denoting the $s \times k$ derivative matrix $\partial \mathbf{g}(\boldsymbol{\beta}_i)/\partial \boldsymbol{\beta}_i{}'$ by $\mathbf{G}(\boldsymbol{\beta}_i)$ and assuming that $\text{rank}[\mathbf{G}(\boldsymbol{\beta}_i)] = s$, the relevant statistic for the test of $\boldsymbol{\eta}_g = \mathbf{0}$ when $\mathbf{b}$ is set *a priori* is given by

$$q_1^* = \left[\mathbf{g}(\mathbf{b}) - \frac{1}{m}\sum_{i=1}^{m}\mathbf{g}(\widehat{\boldsymbol{\beta}}_i)\right]' \widehat{\boldsymbol{\Omega}}_n^{-1} \left[\mathbf{g}(\mathbf{b}) - \frac{1}{m}\sum_{i=1}^{m}\mathbf{g}(\widehat{\boldsymbol{\beta}}_i)\right] \tag{15}$$

where $\widehat{\boldsymbol{\Omega}}_n$ is now defined by

$$\widehat{\boldsymbol{\Omega}}_n = \frac{1}{m^2}\sum_{i,j=1}^{m}\widehat{\mathbf{G}}_i \widehat{\text{Cov}}(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j)\widehat{\mathbf{G}}_j{}', \tag{16}$$

and $\widehat{\mathbf{G}}_i = \mathbf{G}(\widehat{\boldsymbol{\beta}}_)$. (The expression for $\text{Cov}(\widehat{\boldsymbol{\beta}}_i, \widehat{\boldsymbol{\beta}}_j)$ is given in footnote 3.) Then on the null hypothesis of $\boldsymbol{\eta}_g = \mathbf{0}$ (with $\mathbf{b}$ set *a priori*), $q_1^* \overset{a}{\sim} \chi_s^2$.

Turning to the case where $\mathbf{b}$ is defined by (6), Theorem 1 continues to hold with this difference that the appropriate statistic is now given by

$$q_2^* = n^{-1}\widehat{\boldsymbol{\eta}}_g' \widehat{\boldsymbol{\Phi}}_n^{-1} \widehat{\boldsymbol{\eta}}_g \overset{a}{\sim} \chi_s^2, \tag{17}$$

where

$$\widehat{\boldsymbol{\eta}}_g = \mathbf{g}(\widehat{\mathbf{b}}) - \frac{1}{m}\sum_{i=1}^{m}\mathbf{g}(\widehat{\boldsymbol{\beta}}_i), \tag{18}$$

and $\widehat{\boldsymbol{\Phi}}_n$ is defined by (13), although in this more general case $\mathbf{P}_i$ is now given by

$$\mathbf{P}_i = \widehat{\mathbf{G}}_a(\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}' - m^{-1}\widehat{\mathbf{G}}_i(\mathbf{X}_i{}'\mathbf{X}_i)^{-1}\mathbf{X}_i{}', \tag{19}$$

in which $\widehat{\mathbf{G}}_a = \mathbf{G}(\widehat{\mathbf{b}})$ and $\widehat{\mathbf{g}}_i = \mathbf{g}(\widehat{\boldsymbol{\beta}}_i)$. Notice, also that under $\boldsymbol{\eta}_g = \mathbf{0}$, the asymptotic distribution of $q_2^*$ will be a chi-squared with $s(\leq k)$ degrees of freedom. The statistics $q_1^*$ and $q_2^*$ are direct generalisations of $q_1$ and $q_2$ and will reduce to them in the case where $\mathbf{g}(\boldsymbol{\beta}_i) = \boldsymbol{\beta}_i$.

So far, we have limited attention to aggregation bias of the type discussed by Theil (1954) where the bias is defined in terms of the deviations of macro parameters from the simple average of the corresponding micro parameters, as in (4). It is possible that in some circumstances the macro parameters of interest are derived from the micro parameters via a more general function than the average expression $(1/m)\sum \mathbf{g}(\boldsymbol{\beta}_i)$. An obvious example is when the macro parameters are defined as weighted averages of the corresponding micro parameters. To deal with this and other more complicated averaging schemes, we adopt a generalisation of (14) and consider aggregation bias defined as

$$\boldsymbol{\eta}_h = \mathbf{g}(\mathbf{b}) - \mathbf{h}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m) \tag{20}$$

where $\mathbf{h}(\mathbf{b}, \ldots, \mathbf{b}) = \mathbf{g}(\mathbf{b})$. As before, aggregation bias is zero under the micro homogeneity hypothesis, $H_\beta$, but zero aggregation bias (i.e. $\boldsymbol{\eta}_h = \mathbf{0}$) does not necessarily imply $H_\beta$.

The relevant statistics for the test of $\boldsymbol{\eta}_h = \mathbf{0}$ are given by

$$q_1^* = \left[\mathbf{g}(\mathbf{b}) - \mathbf{h}(\widehat{\boldsymbol{\beta}})\right]' \widehat{\boldsymbol{\Omega}}_n^{-1} \left[\mathbf{g}(\mathbf{b}) - \mathbf{h}(\widehat{\boldsymbol{\beta}})\right] \tag{15$'$}$$

and

$$q_2^* = \left[\mathbf{g}(\widehat{\mathbf{b}}) - \mathbf{h}(\widehat{\boldsymbol{\beta}})\right]' \widehat{\boldsymbol{\Phi}}_n^{-1} \left[\mathbf{g}(\widehat{\mathbf{b}}) - \mathbf{h}(\widehat{\boldsymbol{\beta}})\right] \tag{16$'$}$$

where $\mathbf{h}(\boldsymbol{\beta}) = \mathbf{h}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)$. The covariance matrices ($\boldsymbol{\Phi}_n$ and $\boldsymbol{\Omega}_n$ have the same form as before and are given by (13) and (16) respectively, with the difference that the matrix $\widehat{\mathbf{G}}_i$ in (16) need now be replaced by $\widehat{\mathbf{H}_i} = \partial \mathbf{h}(\widehat{\boldsymbol{\beta}})/\partial \boldsymbol{\beta}$, and the matrix $\mathbf{P}_i$ by

$$\mathbf{P}_i = \widehat{\mathbf{G}}_n(\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}' - \widehat{\mathbf{H}}_i(\mathbf{X}_i{}'\mathbf{X}_i)^{-1}\mathbf{X}_i{}'. \tag{21}$$

Once more, $q_1^* \overset{a}{\sim} \chi_s^2$ under the null hypothesis that $\boldsymbol{\eta}_h = \mathbf{0}$, and is the appropriate statistic where $\mathbf{g}(\mathbf{b})$ is given *a priori*. The statistic $q_2^*$ is relevant when $\mathbf{g}(\mathbf{b})$ is estimated from an aggregate equation and is also asymptotically distributed as $\chi_s^2$ under the null hypothesis.

In the application of the above tests to cases where the general functions $\mathbf{g}(\boldsymbol{\beta}_i)$ or $\mathbf{h}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)$ are non-linear in the parameters, special care needs to be exercised in the way the nonlinear restrictions $\boldsymbol{\eta}_g = \mathbf{0}$ or $\boldsymbol{\eta}_h = \mathbf{0}$ are formulated. As has been discussed in the recent literature,[5] when the Wald statistic is used for testing nonlinear restrictions, the value of the test statistic depends on the form of the nonlinear restrictions used in the formulation of the null hypothesis. Although asymptotically this does not matter, in finite samples it is possible to obtain very different values for the Wald statistic by parameterising the hypothesis to be tested in different ways. A simple example which is directly relevant to the empirical application that follows in Section IV helps clarify some of these points. Suppose, for example that we are interested in testing the hypothesis for a single sector $i$ that the long run elasticity of $y_{it}$ with respect to $x_{it}$ in the simple model

$$\log y_{it} = \beta_{i0} + \beta_{i1} \log y_{it-1} + \beta_{i2} \log x_{it} + u_{it}, \tag{22}$$

is equal to, say $c_i$. A usual way of formulating this hypothesis is by means of the nonlinear restriction

$$d_1(\boldsymbol{\beta}_1) = \beta_{i2}/(1 - \beta_{i1}) - c_i = 0. \tag{23}$$

This is not, however, the only way that the hypothesis can be formulated. An alternative and in many ways much more satisfactory formulation of this hypothesis is the linear restriction

$$d_2(\boldsymbol{\beta}_1) = \beta_{i2} + c_1\beta_{i1} - c_i = 0. \tag{24}$$

Although the Wald tests of (23) and (24) are equivalent asymptotically, in small samples, depending on how different $\widehat{c}_i = \widehat{\beta}_{12}/(1 - \widehat{\beta}_{i1})$ is from $c_i$, they can lead to very different results. In this particular example, the linearity of the restriction (24) recommends it over the nonlinear formulation (23),[6] but in general, the choice between alternative parameterisations of nonlinear restrictions is not a straightforward matter.

Similar considerations also apply to our Wald tests of the aggregation bias. Suppose we are interested in testing the hypothesis that the macro long run elasticity of $Y_t = \sum \log y_{it}$ with respect to $X_t = \sum \log x_{it}$ is equal to, say, $c$. When the micro homogeneity hypothesis does not hold, there is no unique method of defining the macro long run elasticity in terms of the micro parameters $\beta_{i1}$ and $\beta_{i2}$. Here we consider two possible methods, the first of which is based on the average of the micro long run elasticities $g(\boldsymbol{\beta}_i) = \beta_{12}/(1 - \beta_{i1}$, namely

$$\epsilon_x^1 = \frac{1}{m}\sum_{i=1}^{m} g(\boldsymbol{\beta}_i) = \frac{1}{m}\sum_{i=1}^{m}\frac{\beta_{i2}}{1 - \beta_{i1}},$$

and the second of which is based on the averages of the micro parameters, namely

$$\epsilon_x^2 = \mathbf{h}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m) = \frac{1}{m}\sum_{i=1}^{m}\beta_{i2}\left(1 - \frac{1}{m}\sum_{i=1}^{m}\beta_{i1}\right)^{-1}.$$

Depending on which of these two definitions are adopted, the null hypothesis of interest can be written as

$$\eta_g = c - \epsilon_x^1 = 0, \tag{25}$$

or

$$\eta_h = c - \epsilon_x^2 = 0. \tag{26}$$

[5] See, for example, Gregory and Veall (1985, 1987), Lafontaine and White (1986), Breusch and Schmidt (1985).

[6] Specifically, in calculating the Wald statistic in the two cases, $\mathrm{Var}(d_2)$ involves the known hypothesised value of $c_i$, while $\mathrm{Var}(\widehat{d}_1)$ involves $\widehat{c}_i$, and so becomes less reliable under $H_0$ as $\widehat{c}_i$ deviates from $c_i$.

As they stand both restrictions are non-linear in the micro parameters and the application of the Wald test to them will be subject to the type of small sample problems emphasised by Gregory and Veall (1985, 1987). Notice, however, that restriction (26) has the advantage that it can be written in linear form:

$$\eta_h' = c - \frac{1}{m}\sum_{i=1}^{m}\beta_{i2} + \frac{1}{m}\sum_{i=1}^{m}c\beta_{i1} = 0, \tag{27}$$

which is the appropriate form to use in the application of the Wald test. Unfortunately, in general the same is not true of the nonlinear restriction (25). The significance of these issues will be illustrated in Section IV.

# III A Misspecification Test of the Disaggregate Model

The tests of aggregation bias advanced above are based on the assumption that the disaggregate model $H_d$ is correctly specified. In particular the tests based on the $q_2$ and $q_2^*$ statistics assume that estimating the macro-parameters directly from the regression of $\mathbf{y}_a$ on $\mathbf{X}_a$, or indirectly by utilising the expression $\sum_{i=1}^{m}\mathbf{C}_i\boldsymbol{\beta}_i$ should not make any difference asymptotically, in the sense that both give consistent estimators of $\mathbf{b}$ under $H_d$. This implication of the disaggregate model can be tested by means of a Durbin-Hausman type misspecification test and suggests basing a test of $H_d$ on the statistic

$$\widehat{\boldsymbol{\eta}}_s = \widehat{\mathbf{b}} - \sum_{i=1}^{m}\widehat{\mathbf{C}}_i\widehat{\boldsymbol{\beta}}_i, \tag{28}$$

where $\widehat{\mathbf{C}}_i$ represents a consistent estimator of $\mathbf{C}_i$ defined by (7).[7] Using the least squares estimates $\widehat{\mathbf{C}}_i = (\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}'\mathbf{X}_i$, $(i = 1, 2, \ldots, m)$, we have

$$\widehat{\boldsymbol{\eta}}_s = (\mathbf{X}_a{}'\mathbf{X}_a)^{-1}\mathbf{X}_a{}'\mathbf{e}_d, \tag{29}$$

where

$$\mathbf{e}_d = \sum_{i=1}^{m}(\mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_i) = \sum_{i=1}^{m}\mathbf{M}_i\mathbf{y}_i, \tag{30}$$

and

$$\mathbf{M}_i = \mathbf{I}_n - \mathbf{X}_i(\mathbf{X}_i{}'\mathbf{X}_i)^{-1}\mathbf{X}_i{}'.$$

Since $(\mathbf{X}_a{}'\mathbf{X}_a)$ is by assumption a non-singular matrix, a test based on $\widehat{\boldsymbol{\eta}}_s$ and $\mathbf{X}_a{}'\mathbf{e}_d$ will be equivalent and for simplicity we use the latter. Suppose now $\mathbf{X}_a$ and $\mathbf{X}_i$ have $p$ variables in common and write[8]

$$\mathbf{X}_a = (\mathbf{X}_{a1}|\mathbf{X}_{a2}); \quad \mathbf{X}_i = (\mathbf{X}_{i1}|\mathbf{X}_{i2}), \quad \text{for all } i,$$

where the $n \times p$ matrix $\mathbf{X}_{a1}$ contains the observations on the common set of variables. It is now easily seen that

$$\mathbf{X}_a{}'\mathbf{e}_d[\underset{p\times 1}{\mathbf{0}} : \underset{(k-p)\times 1}{\mathbf{X}_{a2}{}'\mathbf{e}_d}],$$

and the appropriate statistics on which to base the misspecification test are the non-zero components of $\mathbf{X}_a{}'\mathbf{e}_d$, namely $\mathbf{X}_{a2}{}'\mathbf{e}_d$. Under $H_d$, we have

$$\mathbf{X}_{a2}{}'\mathbf{e}_d = \sum_{i=1}^{m}\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{u}_i, \tag{31}$$

which suggests the following theorem.

**Theorem 2.** *Suppose*

*(i) Assumptions 1–3 hold;*

*(ii) The matrices $n^{-1}(\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{X}_{a2})$ are non-singular in finite samples, and also converge in probability to non-singular matrices;*

[7] See Durbin (1954) and Hausman (1978). Also see Ruud (1984), and Pesaran and Smith (1989) for a unified treatment of misspecification tests in the context of simultaneous equation models.

[8] Examples of such variables include the intercept term, time trends and seasonal dummies.

*(iii) The matrix*

$$\widehat{\mathrm{V}}_n = n^{-1} \sum_{i,j=1}^{m} \widehat{\sigma}_{ij} (\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{M}_j\mathbf{X}_{a2}), \tag{32}$$

*is non-singular for a finite $n$, and converges in probability to the non-singular matrix, $\mathbf{V}$.*

*Then on the hypothesis that the disaggregate model is correctly specified the test statistic*

$$q_3 = n^{-1}\mathbf{e}_d{}'\mathbf{X}_{a2}\widehat{\mathbf{V}}_n^{-1}\mathbf{X}_{a2}{}'\mathbf{e}_d, \tag{33}$$

*is asymptotically distributed as a $\chi^2$ variate with $k-p$ degrees of freedom.*

*Proof.* *See the Mathematical Appendix.*

This theorem complements Theorem 1 and in a sense precedes it. Since Theorem 1 assumes the validity of the disaggregate specification, it is important that the misspecification test of Theorem 2 is carried out before testing for aggregation bias. It is also worth noting that since in general $\sum_{i=1}^m \widehat{\mathbf{C}}_i\widehat{\boldsymbol{\beta}}_i$ is not necessarily a more efficient estimator of $\mathbf{b} = \sum_{i=1}^m \mathbf{C}_i\boldsymbol{\beta}_i$ than $\widehat{\mathbf{b}}$, the familiar Hausman formula for the covariance of $\widehat{\boldsymbol{\eta}}_s$, namely $\mathrm{Cov}(\sum_{i=1}^m \widehat{\mathbf{C}}_i\widehat{\boldsymbol{\beta}}_i)$ is not valid. However, when $\boldsymbol{\beta}_i$ are estimated by the SURE method, the resultant estimators, say$\widetilde{\boldsymbol{\beta}}_i$ will be efficient and the covariance difference formula

$$\mathrm{Cov}(\widetilde{\boldsymbol{\eta}}_s) = \mathrm{Cov}(\widehat{\mathbf{b}}) - \mathrm{Cov}\left(\sum_{i=1}^m \widetilde{\mathbf{C}}_i\widetilde{\boldsymbol{\beta}}_i\right) \geq 0,$$

applies. But even in this case to avoid some of the computational problems that arise because of the possible singularity of $\mathrm{Cov}(\widehat{\mathbf{b}}) - \mathrm{Cov}(\sum_{i=1}^m \widetilde{\mathbf{C}}_i\widetilde{\boldsymbol{\beta}}_i)$, a direct derivation of the variance of $\widetilde{\boldsymbol{\eta}}_s$, along the above lines seems to be more desirable.

# IV An Application

In this section we apply the tests developed in this paper to the annual estimates of aggegate and disaggregate employment demand functions for the U.K. economy presented in PPK. The general log-linear dynamic specification used in the analysis is as follows

$$\begin{aligned} LE_{it} &= \beta_{i1}/m + \beta_{i2}(T_t/m) + \beta_{i3}LE_{i,t-1} + \beta_{i4}LE_{i,t-2} + \beta_{i5}LY_{it} \\ &\quad + \beta_{i6}LY_{i,t-1} + \beta_{i7}LW_{it} + \beta_{i8}LW_{i,t-1} + \beta_{i9}\overline{LY}_{at} \\ &\quad + \beta_{i10}\overline{LY}_{a,t-1} + u_{it}, \quad i = 1,2,3,5,6,\ldots,41 \quad t = 1956, 1957, \ldots, 1984, \end{aligned} \tag{34}$$

where

| | | |
|---|---|---|
| $LE_{it}$ | = | log of man-hours employed in sector $i$ at time $t$; |
| $T_t$ | = | time trend ($T_{1980} = 0$); |
| $LY_{it}$ | = | log of sector $i$ output at time $t$; |
| $LW_{it}$ | = | log of average product real wage rate per man-hours employed in sector $i$ at time $t$; |
| $\overline{LY}_{at}$ | = | average of $LY_{it}$ over the 40 sectors; |
| $m$ | = | number of sectors, ($m = 40$). |

The data cover the whole of the private sector, excluding the Mineral Oil and Natural Gas sector (sector 4) for which the sample size is too short to permit estimation. The rationale behind the above disaggregate model and full details of sources and definitions can be found in PPK.

In order to check the overall validity of the disaggregate specification we first computed the Durbin-Hausman type misspecification test statistic given by (33) in Section III. We obtained a value of 15.9 for this statistic which is distributed as $\chi^2(7)$; this result is just significant at the 5% level and indicates that the disaggregate model may be misspecified. The specification of the disaggregate model requires further consideration, and the following results therefore need to be treated with some caution.

For the purposes of this paper the parameters of interest from the disaggregate model (34) are the long run elasticities with respect to wages and output given respectively by:[9]

$$\epsilon_{iw} = \frac{\beta_{i7} + \beta_{i8}}{1 - \beta_{i3} - \beta_{i4}}, \quad \text{and} \quad \epsilon_{iy} = \frac{\beta_{i5} + \beta_{i6} + \beta_{i9} + \beta_{i10}}{1 - \beta_{i3} - \beta_{i4}}. \tag{35}$$

Table 1: Long Run Elasticities from Restricted Employment Equations* (1956–84)

| Industrial sector | Wage | | Output | |
|---|---|---|---|---|
| 1 Agriculture, forestry and fishing | −0.8981 | (0.2679) | 0.0581 | (0.2904) |
| 2 Coal Mining | −1.9336 | (1.3890) | −1.3866 | (0.4410) |
| 3 Coke | −0.3005 | (0.0418) | 1.6778 | (0.1438) |
| 4 Mineral Oil and Natural Gas | — | — | — | — |
| 5 Petroleum products | −0.6530 | (0.2947) | 0.7560 | (0.3552) |
| 6 Electricity etc. | −0.5379 | (0.4028) | 0.5015 | (0.4090) |
| 7 Public gas supply | −0.2594 | (0.1128) | 1.0311 | (0.5060) |
| 8 Water supply | 0.0 | — | 0.7899 | (0.7082) |
| 9 Minerals and Ores nes. | −0.4870 | (0.2788) | −0.8741 | (0.6483) |
| 10 Iron and steel | −0.7712 | (0.2483) | 2.5657 | (0.8473) |
| 11 Non-ferrous metals | 0.0 | — | 1.9619 | (1.2480) |
| 12 Non-metallic mineral products | −1.4832 | (0.3596) | 2.6847 | (1.1330) |
| 13 Chemicals and manmade fibres | −0.7405 | (0.2146) | 1.5938 | (0.4538) |
| 14 Metal goods nes. | −0.3976 | (0.1987) | 1.0368 | (0.2219) |
| 15 Mechanical engineering | −0.8587 | (0.2457) | 1.2415 | (0.3497) |
| 16 Office machinery etc. | −1.5343 | (2.1250) | 0.0 | — |
| 17 Electrical engineering | −0.0495 | (0.6733) | 1.0277 | (0.8828) |
| 18 Motor vehicles | −0.7238 | (0.4988) | 2.7303 | (1.6720) |
| 19 Aerospace equipment | −0.1763 | (0.1042) | 0.1031 | (0.0953) |
| †20 Ships and other vessels | — | — | — | — |
| 21 Other vehicles | −0.5247 | (0.2978) | 2.1886 | (0.6949) |
| 22 Instrument engineering | −0.5607 | (0.3201) | 0.7715 | (0.3876) |
| 23 Manufactured food | −0.4277 | (0.1530) | 1.7126 | (0.8334) |
| 24 Alcoholic drinks etc. | −0.1302 | (0.3108) | 1.0793 | (0.6737) |
| †25 Tobacco | — | — | — | — |
| 26 Textiles | −0.8320 | (0.2335) | 0.9812 | (0.2690) |
| 27 Clothing and footwear | −0.8101 | (0.1323) | 0.9737 | (0.1454) |
| 28 Timber and furniture | −0.1700 | (0.1214) | 0.6627 | (0.1240) |
| 29 Paper and board | −0.3938 | (0.0968) | 0.9856 | (0.2567) |
| 30 Books etc. | 0.2074 | (0.2758) | 0.1818 | (0.1955) |
| 3i Rubber and plastic products | −0.5767 | (0.3846) | 1.2662 | (0.5789) |
| 32 Other manufactures | 0.0 | — | 0.5903 | (0.2431) |
| 33 Construction | −0.6453 | (1.0200) | 0.5872 | (0.9331) |
| 34 Distribution etc. | −0.6965 | (0.3259) | −.01248 | (0.1323) |
| 35 Hotels and catering | −0.6602 | (0.4499) | 1.2205 | (0.7663) |
| 36 Rail transport | −0.3735 | (0.3632) | 2.0843 | (0.7845) |
| 37 Other land transport | 0.0 | — | 0.4203 | (0.1983) |
| 38 Sea, air and other | −0.2354 | (0.1985) | 0.5309 | (0.4275) |
| 39 Communications | −0.0267 | (0.1905) | 0.9905 | (0.6643) |
| 40 Business services | 0.0 | — | 0.2333 | (0.1294) |
| 41 Miscellaneous services | −0.8108 | (0.7865) | 1.2228 | (1.0810) |
| Mean of long run elasticities | −0.5233 | | 0.9489 | |
| Standard deviation of elasticities | 0.4437 | | 0.8867 | |
| Median of long run elasticities | −0.5247 | | 0.9812 | |

* The estimates reported in this table are based on the results in table 2 of PPK. The bracketed figures are the estimated standard errors.

† These industries are excluded from the analysis. See the text for further explanation.

Table 1 presents estimates for these elasticities derived from the set of restricted disaggregate employment equations estimated by PPK (table 2), with asymptotically valid standard errors in parentheses.[10] For two sectors, 20 and 25, the employment equations estimated by PPK do not possess long run solutions so that there are no corresponding elasticities in Table 1. These two industries are excluded from the subsequent analysis. For a few sectors, PPK found no significant response with respect to the real wage variable, and in one sector no response was found with respect to the output variable. In these cases the estimates of the long run elasticity in the table are set equal to zero and no standard errors are given. The last three rows of Table 1 present the mean, standard deviation and the median of the distribution of the estimates of the elasticities across the sectors. In the case of both sets of elasticity estimates the mean and median are approximately equal showing that the distributions are close to being symmetric. Both of the standard deviations are large highlighting the considerable variation in the employment responses between sectors. This in itself can be viewed as an argument for the use of disaggregated analysis. We now consider the application of the tests of aggregation bias developed in Section II (namely the $q_1^*$ and the $q_2^*$ tests) to the disaggregated long run elasticities of Table 1 and those of the corresponding restricted aggregate equation given by:[11]

$$\begin{aligned} LE_{at} = {} & -140.93 + 0.6935 LE_{a,t-1} + 0.4665 LY_{at} - 0.3948 LW_{at} + \widehat{u}_{at}, \qquad (36)\\ & (17.177)\ (0.0424) \qquad\qquad (0.0488) \qquad\quad (0.0387) \\ & \overline{R}^2 = 0.9954, \quad \widehat{\sigma}^2 = 0.3666, \quad n = 29\ (1954\text{–}84), \\ & \chi^2_{SC}(1) = 1.28, \quad \chi^2_{FF}(1) = 2.45, \quad \chi^2_N(2) = 4.41, \quad \chi^2_H(1) = 3.45, \end{aligned}$$

where $LE_{at}$, $LW_{at}$ and $LY_{at}$ are the sums of $LE_{it}$, $LW_{it}$, and $LY_{it}$ over the 38 sectors respectively. $\overline{R}^2$ is the adjusted $R^2$, $\widehat{\sigma}^2$ is the estimated standard error of the regression, and $\chi^2_{SC}$, $\chi^2_{FF}$, $\chi^2_N$ and $\chi^2_H$ are respectively the chi-squared statistics for residual serial correlation, functional form misspecification, normality, and homoskedasticity of the disturbances.[12] The estimates of the long run real wage and output elasticities based on (36) are $-1.2880$ (0.2947) and 1.5221 (0.3386) respectively. The numbers in parentheses are asymptotically valid standard errors. It is clear that these results are consistent with the hypothesis of wage and output elasticities of $-1$ and $+1$ respectively. The relevant statistic for testing the hypothesis that the average of the disaggregate elasticities of Table 1 is equal to unity is given by $q_1^*$, (15). In this case $\mathbf{g}(\mathbf{b}) = --1$, $\mathbf{g}(\widehat{\boldsymbol{\beta}}_i) = \widehat{\epsilon}_{iw}$ for the real wage variable and $\mathbf{g}(\mathbf{b}) = 1$, $\mathbf{g}(\widehat{\boldsymbol{\beta}}_i) = \widehat{\epsilon}_{iy}$ for the output variable where the long run elasticities $\epsilon_{iw}$ and $\epsilon_{iy}$ are already defined in (35). The hypothesis of a unit average long run output elasticity can not be rejected even at the 10% level, since in this case $q_1^*$ equals 0.104 based on an estimated value for the average disaggregate output elasticity of 0.9489. In contrast, the value of $-0.5233$ obtained for the average disaggregate wage elasticity is significantly different from $-1$ with $q_1^*$ taking the value of 19.27 in this case.[13] The estimated $q_2^*$ statistics reinforce the finding that the aggregate and disaggregate results differ significantly. The values of this statistic for the wage and output elasticities are 10.95 and 4.97 respectively, rejecting the null hypothesis of no aggregation bias in both cases.[14]

We also considered the alternative aggregate restrictions involving the responsiveness of employment to real wage and output changes corresponding to (26) in the simple model of SectionII. As noted there, these restrictions can also be written in a linear form as in (27) and the $q_1^*$ statistic given by $(15)'$ was computed here using both linear and nonlinear forms of the restrictions. This allows us to examine the practical importance of the issue of parameterisation of the nonlinear restrictions in the case of the Wald tests discussed in Section II. For wage responsiveness, the values of the $q_1^*$ statistic were 166.88 and 66.01 for the restriction forms (26) and (27) respectively, both forms of the test massively rejecting the null hypothesis of a long run real wage elasticity of minus unity. For output responsiveness, the values of the $q_1^*$ statistic were 0.129 and 0.126 respectively, so that the null hypothesis of a unit long run output

---

[9] The formula for the output elasticity allows for the long run effect of the sectoral changes on employment of the $i$th sector both directly through the terms $LY_{it}$ and $LY_{i,t-1}$, and indirectly through the aggregate output effects $\overline{LY}_{at}$ and $\overline{LY}_{a,t-1}$.

[10] Details of the exclusion restrictions imposed for each sector and the diagnostic statistics computed for each equation can also be found in PPK.

[11] This result corresponds to equation (7.4) in PPK, reestimated to exclude sectors 20 and 25.

[12] See Pesaran and Pesaran (1987) for further details and the relevant algorithms.

[13] In the present application, $q_1^*$ and $q_2^*$ are both distributed asymptotically as $\chi^2_1$ under the null hypothesis.

[14] The $q_2^*$ statistics are calculated using (17) replacing $\mathbf{g}(\widehat{\mathbf{b}})$ and $\mathbf{g}(\widehat{\boldsymbol{\beta}}_i)$ by their corresponding aggregate and disaggregate long run elasticity estimates.

elasticity cannot be rejected in the case of either formulation.[15] Clearly, the alternative parameterisations considered here have a considerable effect on the value of the statistic obtained for the wage restriction although the results are unaffected qualitatively.

In conclusion, our estimates of the disaggregate labour demand relationships show that there is considerable variation across sectors so that important information may be lost in working with aggregate figures. This is confirmed by the application of the tests developed in the paper. Significant aggregation biases are found in the estimates of a variety of measures of the responsiveness of employment to real wage and output changes based on our aggregate and disaggregate employment equations. The problem of aggregation bias seems, however, to be much more serious for the estimates of the long run real wage elasticity as compared to the estimates of the long run output elasticity.

# A Mathematical Appendix

*Proof of Theorem 1.* Under $H_0$ defined by (8), the statistic $q_2$ in (10) can be written as

$$q_2 = \mathbf{d}_n{}'\mathbf{d}_n, \tag{A.1}$$

where

$$\mathbf{d}_n = \sum_{i=1}^{m} \mathbf{z}_{in}, \tag{A.2}$$

and

$$\mathbf{z}_{in} = n^{-\frac{1}{2}}\widehat{\mathbf{\Phi}}_n^{-1}\mathbf{P}_i\mathbf{u}_i, \quad i = 1, 2, \ldots, m. \tag{A.3}$$

The matrices $\widehat{\mathbf{\Phi}}_n$ and $\mathbf{P}_i$ are defined by (13) and (11) in the text, respectively. The proof we offer here has two stages: we first show that for each $i$ and for any real $k \times 1$ vector $\boldsymbol{\lambda}$ such that $\boldsymbol{\lambda}'\boldsymbol{\lambda} = 1$, $\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} \mathrm{N}(0, \phi_{ii})$ where $\phi_{ii} > 0$. Using this result in (A.2), we then show that $\boldsymbol{\lambda}'\mathbf{d}_n \overset{a}{\sim} \boldsymbol{\lambda}'\mathbf{d}$, where $\mathbf{d} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_k)$. From this it follows that $\mathbf{d}_n \overset{a}{\sim} \mathrm{N}(\mathbf{0}, \mathbf{I}_k)$, and $\mathbf{d}_n{}'\mathbf{d}_n \overset{a}{\sim} \chi^2_k$. See proposition 5.1 in White (1984).

Under Assumptions 1–3 it readily follows that

$$\operatorname*{plim}_{n\to\infty}(\widehat{\sigma}_{ij}) = \sigma_{ij}, \quad \operatorname*{plim}_{n\to\infty}(\mathbf{\Phi}_n) = \mathbf{\Phi},$$

where

$$\mathbf{\Phi} = \sum_{i,j=1}^{m} \sigma_{ij}\mathbf{Q}_{ij},$$

and the matrices $\mathbf{Q}_{ij}$ defined by

$$\begin{aligned}\mathbf{Q}_{ij} &= \operatorname*{plim}_{n\to\infty}(n^{-1}\mathbf{P}_i\mathbf{P}_j{}') \\ &= \mathbf{\Sigma}_{aa} - \frac{1}{m}\mathbf{\Sigma}_{aa}^{-1}\mathbf{\Sigma}_{aj}\mathbf{\Sigma}_{jj}^{-1} - \frac{1}{m}\mathbf{\Sigma}_{ii}^{-1}\mathbf{\Sigma}_{ia}\mathbf{\Sigma}_{aa}^{-1} + \frac{1}{m^2}\mathbf{\Sigma}_{ii}^{-1}\mathbf{\Sigma}_{ij}\mathbf{\Sigma}_{jj}^{-1},\end{aligned}$$

are finite for all $i$ and $j$ and are non-singular for $i = j$. Now noting that by assumption $\mathbf{\Phi}$ is also non-singular we have

$$\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} n^{-\frac{1}{2}}\boldsymbol{\mu}'\mathbf{P}_i\mathbf{u}_i = n^{-\frac{1}{2}}\sum_{t=1}^{n}\delta_{it}u_{it}, \tag{A.4}$$

where $\boldsymbol{\mu} = \mathbf{\Phi}^{-1}\boldsymbol{\lambda}$, and $\delta_{it}$ stands for a typical element of vector $\mathbf{P}_i{}'\boldsymbol{\mu}$. It is now easily seen that under assumptions of the theorem, the conditions for the application of the version of Liapounov's Theorem cited in (White, 1984, theorem 5.10) to the right hand side of (A.4), which is a sum of independently, but non-identically distributed random variables, are met and

$$\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} \mathrm{N}(0, \phi_{ii}), \quad \text{where} \quad \phi_{ii} = \sigma_{ii}\boldsymbol{\mu}'\mathbf{Q}_{ii}\boldsymbol{\mu} > 0.$$

[15] The estimates of the aggregate long run elasticities of output and real wages underlying the restriction form (26) are 0.9770 and $-0-4551$ respectively, as compared to the estimates 0.9909 and $-0.7850$ underlying the restriction form (27).

Therefore, asymptotically $\boldsymbol{\lambda}'\mathbf{d}_n = \sum_{i=1}^m \boldsymbol{\lambda}'\mathbf{z}_{in}$ is distributed as a linear function of $m$ normal variates and itself will be distributed normally with zero mean and variance[16]

$$\lim_{n\to\infty} \mathrm{V}(\boldsymbol{\lambda}'\mathbf{d}_n) = \boldsymbol{\lambda}'\boldsymbol{\lambda} = 1.$$

Hence, for a finite $m$, $\mathbf{d}_n \overset{a}{\sim} \mathrm{N}(\mathbf{0}, \mathbf{I}_k)$, and $\mathbf{d}_n{}'\mathbf{d}_n \overset{a}{\sim} \chi^2_k$. **Q.E.D.**

*Proof of Theorem 2.* The proof is similar to that presented for Theorem 1. Under $H_d$ the statistic $q_3$ defined by (33) can be written as $q_3 = \mathbf{d}_n{}'\mathbf{d}_n$ where $\mathbf{d}_n$ is defined by (A.2), but $\mathbf{z}_{in}$ is now given by $z_{in} = n^{-\frac{1}{2}}\widehat{\mathbf{V}}_n^{-\frac{1}{2}}\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{u}_i$. Since by assumption $\widehat{\mathbf{V}}_n$ converges in probability to a non-singular matrix, say $\mathbf{V}$, we also have

$$\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} n^{-\frac{1}{2}}\boldsymbol{\lambda}\mathbf{V}^{-\frac{1}{2}}\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{u}_i = n^{-\frac{1}{2}}\boldsymbol{\mu}'\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{u}_i,$$

where $\boldsymbol{\lambda}$ is now a $(k-p)\times 1$ vector of constants such that $\boldsymbol{\lambda}'\boldsymbol{\lambda} = 1$, and $\boldsymbol{\mu} = \mathbf{V}^{-\frac{1}{2}}\boldsymbol{\lambda}$. Denoting the $t$th element of $\mathbf{M}_i\mathbf{X}_{a2}\boldsymbol{\mu}$ by $\eta_{it}$ we now have

$$\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} n^{-frac12}\sum_{t=1}^{n}\eta_{it}u_{it}, \tag{A.5}$$

which is a sum of independently, but non-identically distributed random variables. As in the proof of Theorem 1, it is easily seen that under assumptions of Theorem 2, the Liapounov's theorem ((White, 1984, theorem 5.10)) is applicable to (A.5) and

$$\boldsymbol{\lambda}'\mathbf{z}_{in} \overset{a}{\sim} \mathrm{N}(0, \psi_{ii}),$$

where

$$\psi_{ii} = \boldsymbol{\mu}'\left[\operatorname*{plim}_{n\to\infty}(n^{-1}\mathbf{X}_{a2}{}'\mathbf{M}_i\mathbf{X}_{a2})\right]\boldsymbol{\mu} > 0.$$

Hence, by a similar reasoning as in the proof of Theorem 1, we have

$$\boldsymbol{\lambda}'\mathbf{d}_n \overset{a}{\sim} \mathrm{N}(0,1), \quad \text{and} \quad \mathbf{d}_n \overset{a}{\sim} N(bfzero, \mathbf{I}_{k-p}),$$

which establishes that

$$q_3 = \mathbf{d}_n{}'\mathbf{d}_n \overset{a}{\sim} \chi^2_{k-p}.$$

**Q.E.D.**

## References


Boot, J. C. G. and G. M. de Wit (1960). Investment demand: an empirical contribution to the aggregation problem. *International Economic Review 1*, 3–30.

Breusch, T. and P. Schmidt (1985). *Alternative forms of the Wald test: how long is a piece of string?* Unpublished manuscript.

Durbin, J. (1954). Errors in variables. *Review of International Statistical Institute 22*, 23–32.

Gregory, A. W. and M. R. Veall (1985). Formulating Wald tests of nonlinear restrictions. *Econometrica 53*, 1465–1468.

Gregory, A. W. and M. R. Veall (1987). Formulating Wald tests of the restrictions implied by the rational expectations hypothesis. *Journal of Applied Econometrics 2*, 61–68.

Grunfeld, Y. and Z. Griliches (1960). Is aggregation necessarily bad? *Review of Economics and Statistics 42*, 1–13.

Gupta, K. L. (1971). Aggregate bias in linear economic models. *International Economic Review 12*, 293–305.


[16] Notice that since $\lim_{n\to\infty}\mathrm{V}(\boldsymbol{\lambda}'\mathbf{z}_{in}\mathbf{z}_{jn}{}'\boldsymbol{\lambda}) = \sigma_{ij}\boldsymbol{\mu}'\mathbf{Q}_{ij}\boldsymbol{\mu} = \phi_{ij}$, then $\lim_{n\to\infty}\mathrm{V}(\boldsymbol{\lambda}'\mathbf{d}_n) = \lim_{n\to\infty}\mathrm{V}(\sum_{i=1}^m \boldsymbol{\lambda}'\mathbf{z}_{in}) = \sum_{i,j=1}^m \phi_{ij} = \boldsymbol{\mu}'(\sum_{i,j=1}^m \sigma_{ij}\mathbf{Q}_{ij})\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\mu} = \boldsymbol{\lambda}'\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{-\frac{1}{2}}\boldsymbol{\lambda} = 1$.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*, 1251–1271.

Lafontaine, F. and K. J. White (1986). Obtaining any Wald statistic you want. *Economics Letters 21*, 35–40.

Lovell, C. A. K. (1973). A note on aggregation bias and loss. *Journal of Econometrics 1*, 301–311.

Orcutt, G. H., H. W. Watts, and J. B. Edwards (1968). Data aggregation and information loss. *American Economic Review 58*, 773–787.

Pesaran, M. H. and B. Pesaran (1987). *Microfit: An Interactive Econometric Software Package*. Oxford: Oxford University Press.

Pesaran, M. H., R. G. Pierse, and M. S. Kumar (1989). Econometric analysis of aggregation in the context of linear prediction models. *Econometrica 57*, 861–888.

Pesaran, M. H. and R. J. Smith (1989). A unified approach to estimation and orthogonality tests in linear single equation econometric models. *Journal of Econometrics*. (forthcoming).

Ruud, P. A. (1984). Tests of specification in econometrics. *Econometric Reviews 3*, 211–242.

Sasaki, K. (1978). An empirical analysis of linear aggregation problems. the case of investment behavior in Japanese firms. *Journal of Econometrics 7*, 313–331.

Theil, H. (1954). *Linear Aggregation of Economic Relations*. Amsterdam: North-Holland.

White, H. (1984). *Asymptotic Theory for Econometricians*. Orlando, FL: Academic Press.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association 57*, 348–368.