THE UNIVERSITY of York

Discussion Papers in Economics

No. 2000/50

The Demand for Education and the Production of Local Public Goods

by

David Mayston

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

THE DEMAND FOR EDUCATION

AND THE

PRODUCTION OF LOCAL PUBLIC GOODS

DAVID J. MAYSTON*

University of York

*The author is Professor of Public Sector Economics, Finance and Accountancy, and Director of the Centre for Performance Evaluation and Resource Management, at the University of York.

ABSTRACT

Improving the educational outcomes which schools achieve in the primary and secondary education sectors has become a central public policy goal, to which large sums of public money have been devoted. Being able to estimate the educational production function between pupil educational achievements, resource inputs and characteristics of the pupil intake in an unbiased way can help progress both ***ex ante*** policy formation and ***ex post*** effectiveness monitoring. Such an unbiased estimation requires recognition of not only the ***supply-side*** concept of the educational production function but also several ***demand-side*** relationships affecting the demand for school places, the socio-economic characteristics of a school's pupil intake, the quality of teaching staff a school can recruit, and the determination of local property prices. Failure to recognise these additional inter-relationships through the use of standard single-equation Ordinary Least Squares multivariate regression will result in ***multiple*** sources of ***cumulative downward bias*** in the estimated importance of resource variables in influencing pupil educational outcomes, in ways which are analysed in this paper. The analysis of this paper calls into question the conclusions drawn by Hanushek and others, from many of the earlier statistical studies of the educational production function, of there existing no substantial link between educational resourcing and educational outcomes.

1. INTRODUCTION

Education has many of the economic characteristics of a local public good, being typically provided by local schools to a local population of the school's pupil intake. Education is produced at different levels by different individual schools, with these different levels of production impacting upon the educational outcomes which the pupils of each school achieve. Improving the educational outcomes which schools achieve in the primary and secondary education sectors has become a central public policy goal, to which large sums of public money have been devoted in return for Public Service Agreements on the target levels of educational outcomes which are expected to be achieved by the education system as a whole in return for increased levels of public funding (see Mayston, 2000a).

Being able to correctly identify the levels of educational outcomes which are achievable from different levels of resource inputs, for given characteristics of the pupil intake, and the levels of resourcing which are required to achieve any given desired level of educational outcomes, is of importance not only for public expenditure planning at the national level. It is also important for the determination of the levels of resources which should be allocated to different individual schools with differing characteristics of their pupil intake to achieve individual target levels of educational outcomes, and for monitoring the levels of educational outcomes which can be expected from these individual schools from their available resource budgets.

In the case of a single scalar measure of educational outcomes, q, for any given school, such as the school's average GCSE point scores for its pupils, we may seek to represent the relationship between the levels of educational outcomes, q, that are achievable from different levels of resource inputs and different characteristics of the school's pupil intake, through the concept of an ***educational production function***, f, of the form:

$$q = f(r, s) \tag{1.1}$$

where r is a vector of school resource inputs. s is here a vector of characteristics of the school's pupil intake, which can include variables relating to the pupils' home and parental circumstances which may influence their educational outcomes for given levels of school resourcing.In the case of a vector of educational outcomes, q, for the given school, we may write the production function in the implicit multi-product form:

$$g(q, r, s) = 0 \tag{1.2}$$

with g specifying the different combinations of the individual educational outputs within the school output vector q that are achievable from different given levels of the school resource input vector r and of the vector of characteristics of the school's pupil intake s.

The positive role of the resource vector, r, in improving educational outcomes has been called into question by Hanushek (1979, 1986), based upon a review of earlier statistical studies which allegedly find no statistically significant relationship between school expenditure and pupil outcomes. The method of statistical estimation used in these earlier studies has been predominantly that of Ordinary Least Squares (OLS) multivariate regression analysis. However, whether or not this estimation technique provides unbiased estimates of the parameters of the relationship being investigated depends upon whether or not the assumptions which the OLS technique makes are fully satisfied (see Johnston, 1984; Gujarati, 1995). One of the main prior assumptions made by OLS is that the disturbance term in the multivariate regression equation is uncorrelated with the explanatory variables that are used in the regression equation. A failure to satisfy this prior assumption will in general result in the parameter estimates which OLS produces being ***biased*** away from their true underlying values, even if a large sample of data is used for the estimation. In this paper, we will investigate a number of important sources of this bias, in the context of the estimation of the educational production function.

2. THE ENDOGENEITY PROBLEM

For the sake of concreteness, we will first examine an efficient school with a single educational output, q, and a Cobb-Douglas production function of the form:

$$q_i = A.T_i^{\hat{a}_2}.Q_i^{\hat{a}_3}.N_i^{\hat{a}_4}.K_i^{\hat{a}_5}.m_i^{\hat{a}_6}.S_i^{\hat{a}_7} \qquad (2.1)$$

where the i subscripts denote the corresponding variables for school i.

The variable T_i denotes the school's teacher-pupil ratio, whilst Q_i is an index of teacher quality, the variable N_i represents non-teaching staff per pupil, K_i represents non-staff inputs per pupil, m_i denotes total pupil numbers, and the variable S_i is a measure of the socio-economic background of the pupil intake. Higher levels of S_i are assumed here to indicate a more advantaged pupil background. Whilst S_i is here a scalar variable, the analysis may be readily extended to include a whole set of different socio-economic variables that characterise the pupil intake and which may influence the school's educational outcomes for a given set of resource inputs. In any given period of time, the parameter A is a constant, but may increase over time if technological change makes possible higher levels of productivity within schools. The parameters $\hat{a}_2$, ..., $\hat{a}_7$ correspond to the proportional change in educational output which can be achieved by a unit proportional change in the input of each of the corresponding variables, with $\hat{a}_6$ indicating the extent of the economies of scale with respect to pupil numbers which exist within the educational production function at the school level.

For any individual school i, whether fully efficient or not, we will write:

$$q_i \;=\; \mathring{a}_i.A.T_i^{\hat{a}_2}.Q_i^{\hat{a}_3}.N_i^{\hat{a}_4}.K_i^{\hat{a}_5}.m_i^{\hat{a}_6}.S_i^{\hat{a}_7} \tag{2.2}$$

where $\mathring{a}_i$ denotes an index of efficiency of school i, yielding the linear regression equation:

$$\log\ q_i \;=\; \hat{a}_0 + \hat{a}_2.\log T_i + \hat{a}_3.\log Q_i + \hat{a}_4.\log N_i + \hat{a}_5.\log K_i + \hat{a}_6.\log m_i + \hat{a}_7.\log S_i + \log\ \mathring{a}_i \tag{2.3}$$

when the variables are expressed in logarithmic form. Equation (2.3) may be re-written in the form:

$$b_{01} + b_{11}\log\ q_i + b_{21}.\log T_i + b_{31}.\log Q_i + b_{41}.\log N_i + b_{51}.\log K_i + b_{61}.\log m_i + b_{71}.\log S_i \;=\; u_{i1} \qquad (2.3a)$$

where $b_{01} \equiv -\hat{a}_0$, $b_{11} \equiv 1$, $b_{k1} \equiv -\hat{a}_k < 0$ for k = 2,..., 7 if S_i denotes a measure of socio-economic advantage, and $u_{i1} \equiv \log\ \mathring{a}_i$.

Under the standard assumption of OLS that the disturbance term is of zero mean and positive variance, log $\mathring{a}_i$ in (2.3) can take on positive values for some schools in the sample. One main problem with the use of OLS to estimate the underlying production function is then that OLS will estimate a ***statistically average*** 'educational production function', rather than the fully efficient educational production function, corresponding to the underlying education production possibility ***frontier***, in which log $\mathring{a}_i$ would take on only non-positive values. A positive relationship may then exist between resource inputs and educational outcomes along the educational production frontier, even if no such positive relationship exists within the estimated statistically average 'educational production function'.

However, even if we ignore this problem, important additional considerations arise when a second main assumption of OLS is broken. This is the assumption that the explanatory variables are uncorrelated with the random disturbance term in an equation such as (2.3). As noted above, the parameter estimates which OLS produces may then be ***biased estimates*** of the underlying parameters, such as $\hat{a}_2$, ..., $\hat{a}_7$ in equation (2.1), that ***deviate from their true underlying values*** , even if an infinitely large sample of observations is analysed, with the parameter estimates which OLS producing of the influence of each of the school inputs on the school's educational output then failing to be ***consistent estimators*** of their true underlying value (see Johnston, 1984; Gujarati, 1995).

One way in which non-zero correlations may be generated between the disturbance term and the explanatory variables is if there are additional inter-relationships between the variables beyond those defined by the ***supply-side*** concept of the educational production function. In order to form a complete model of these inter-relationships, we will add further equations to the single equation that is involved in (2.1a). These additional equations will recognise that there are likely to be additional ***demand-side*** inter-relationships between the different variables beyond those that are involved in the supply-side concept of the educational production function in (2.3) or (2.3a). The importance of educational demand-side relationships, and their relationship to the optimal allocation of resources within schools, is stressed in Mayston (1996).

Equation (2.3), or equivalently equation (2.3a), will be the first equation in our overall model of the inter-relationships which may exist between the different variables of pupil outcomes, resource input and characteristics of the pupil intake. Within this model, the coefficient b_{kj} will be used to designate that on the kth variable in the jth equation of our model. Our first variable here is log q_i , the log of examination scores, our second variable is log T_i , the log of the school's teacher-pupil ratio, our third variable is log Q_i, the log of teacher quality, and so on. b_{21} is then, for example, the coefficient on the second variable in our first equation (2.3a).

3. INPUT INTER-RELATIONSHIPS

The first such inter-relationship involves the teacher-pupil ratio, T_i , for school i. This will depend upon a number of factors. Under the devolved budgetary arrangements of Local Management of Schools, or the current *Fair Funding* framework (DfEE, 1998), these are likely to include the level of the school's available income per pupil, $\div_i$, which will form the eighth variable in our model. They are also likely to include the local cost, p_{Ti} , of attracting teaching staff of a given quality to the local area, the local price, p_{Ni} , of non-teaching staff and the local cost, p_{Ki} , of other inputs that must be purchased out of the school's total expenditure budget, and the quality, Q_i, of the teaching staff which the school is successful in recruiting. These factors will then enter into the demand function of the individual school for teaching staff, which we may write in the form:

$$\log T_i \;=\; a_{02} \;+ a_{32} \,.\, \log Q_i \;+ a_{82} \,.\log \div_i \;+ a_{92} \,.\, \log p_{Ti} + a_{10,2} \,.\, \log p_{Ni} + a_{11,2} \,.\log p_{Ki} + u_{i\,2} \tag{3.1}$$

where we expect $a_{82} > 0$, i.e. the teacher-pupil ratio to be an increasing function of the school's income per pupil, and $a_{92} < 0$, i.e.the teacher-pupil ratio to be a decreasing function of the price it must pay for teachers, given its total budget. $a_{10,2}$ and $a_{11,2}$ may be positive, negative or zero, depending upon the relative strength of the income and substitution effects that a rise in the price of these other inputs produces for the demand for teachers within the school. We would expect $a_{32} < 0$, so that the school faces a choice within its total budget between hiring a larger quantity of teachers of a lower quality and hiring fewer teachers of higher quality who are more in demand elsewhere, and who can command higher salaries. The random disturbance term, $u_{i\,2}$, for school i reflects the significant degree of variation which may still exist across individual schools in their teacher-pupil ratios even after the above systematic factors are taken into account (see e.g. Audit Commission, 1993 - 1996).

(3.1) may be re-written in the form:

$$b_{02} \;+ b_{22} \,.\, \log T_i + b_{32} \,.\, \log Q_i \;+ b_{82} \,.\log \div_i \;+ b_{92} \,.\, \log p_{Ti} + b_{10,2} \,.\, \log p_{Ni} + b_{11,2} \,.\log p_{Ki} = u_{i\,2} \tag{3.1a}$$

where $b_{02} \equiv - a_{02}$, $b_{22} = 1$, $b_{32} \equiv - a_{32} > 0$, $b_{82} \equiv - a_{82} < 0$, $b_{92} \equiv - a_{92} > 0$, $b_{10,2} \equiv - a_{10,2}$, $b_{11,2} \equiv - a_{11,2}$. Again we adopt the convention that the coefficient $b_{k\,j}$ is used to designate the coefficient on the kth variable in the j th equation of our model, so that b_{32} is here the coefficient in this second equation on our third variable, $\log Q_i$. The fact that not all variables (such as $\log q_i$ here) are entering each of the equations of our model imposes a structure on the model which will assist in overcoming problems of ***identifiability*** of the education production function within our simultaneous equations model.

The quality of teaching staff which the school attracts plays a potentially significant role as a school resource input in the educational production function (2.1). The importance of teacher quality in boosting the educational output of schools has been stressed, for instance, by Murdane (1996). It is important to note that teacher quality is also likely to be a significant additional source of ***endogeneity*** in the inter-relationships between the variables involved in (2.3). The ability of an individual school i to attract teaching staff of high quality is itself likely to depend upon a number of variables. These may include the level of the school's examination success, q_i, the nature of its pupil intake, S_i , the size of the school, m_i, and its associated scope for specialised teaching and supporting facilities. They may also include its teacher-pupil ratio, T_i , and the level of its support from non-teaching staff, N_i , and non-staffing resources per pupil, K_i . Schools which are well endowed in these directions may be better able to attract higher quality teachers than those schools without these advantages. This itself will, however, establish an additional inter-relationship between the relevant variables.

If, other things being equal, higher quality teachers are attracted more to schools with a higher levels of examination performance, q_i , a positive correlation will exist between the explanatory variable of teacher quality, Q_i , in the multiple regression equation (2.3) and q_i . Since the level of q_i in (2.3) increases with the value of the disturbance term logε_i , this inter-relationship will tend to generate a positive correlation between the explanatory variable of teacher quality, Q_i , in the multiple regression equation (2.3) and the disturbance term logε_i . As noted above, such a non-zero correlation breaches one of the main assumptions that must be hold for the valid application of the standard OLS estimation procedure. This breach will bias the parameter estimates of the impact of Q_i in the educational production function (2.1) away from its true value.

The quality of teaching staff which a school i attracts may also depend upon the level of pay, p_{Ti} , for teachers which is available to its staff. The cost of local housing, as influenced by local house prices, p_{Hi} , may well also influence the quality of staff which school i can recruit and retain. Schools in central London may then have more difficulty in attracting high quality staff because of high local house prices for houses of a given quality. If these schools also have a high intake of disadvantaged and disruptive pupils, the influence of S_i on Q_i may further increase the difficulty of inner city areas in attracting high quality staff. Even if not all teaching staff live close to the school, high local house prices may necessitate their commuting from a greater distance to teach in the school, reducing its relative attractions as a school in which to teach. The above considerations imply an inter-relationship between teacher quality, Q_i , and the other variables of the form:

$$\log Q_i = a_{03} + a_{13} .\log q_i + a_{23} .\log T_i + a_{43} .\log N_i + a_{53}. \log K_i + a_{63} . \log m_i$$
$$+ a_{73}.\log S_i + a_{93} .\log p_{Ti} + a_{12,3}. \log p_{Hi} + u_{i3} \tag{3.2}$$

where u_{i3} is a random disturbance term. We would expect that $a_{13} > 0$, so that schools with superior examination performance and a high rating in school performance tables are better able to attract higher quality teachers. We would also expect that $a_{23} > 0$, $a_{43} > 0$, and $a_{53} > 0$, so that schools with more favourable teacher-pupil ratios and other supporting resources are better able to attract higher quality teachers. Larger schools that are able to offer more specialised teaching may also be more able to attract better quality teachers, implying $a_{63} > 0$. If there is a statistical association between an advantaged socio-economic background, as measured by S_i , and the proportion of pupils who are disruptive, then we may also have $a_{73} > 0$, so that schools in tougher inner-city areas may have more difficulty in attracting higher quality staff. If higher local house prices also make it more difficult to attract higher quality staff, then we would expect $a_{12,3} < 0$. Higher local salaries for teachers will, however, make it easier to attract higher quality teachers, with $a_{93} > 0$.

Equation (3.2) may be written in the form:

$$b_{03} + b_{13} .\log q_i + b_{23} .\log T_i + b_{33}. \log Q_i + b_{43} .\log N_i + b_{53}. \log K_i + b_{63} . \log m_i$$
$$+ b_{73}.\log S_i + b_{93} .\log p_{Ti} + b_{12,3}. \log p_{Hi} = u_{i3} \qquad (3.2a)$$

where $b_{03} \equiv - a_{03}$, $b_{13} \equiv - a_{13} < 0$, $b_{23} \equiv - a_{23} < 0$, $b_{33} = 1$, $b_{43} \equiv - a_{43} < 0$, $b_{53} \equiv - a_{53} < 0$, $b_{63} \equiv - a_{63} <0$, $b_{73} \equiv - a_{73} <0$, $b_{93} \equiv - a_{93} < 0$, $b_{12,3} \equiv - a_{12,3} > 0$.

The demand of school i for non-teaching staff per pupil, N_i, is likely to depend upon its available income per pupil, $\div_i$, upon the local price, p_{Ni} , of non-teaching staff , and on the prices which it must pay for teaching staff and non-staff expenditure out of its overall expenditure per pupil budget. In log-linear form, this implies an inter-relationship between N_i and the other variables of the form:

$$\log N_i = a_{04} + a_{84} .\log \div_i + a_{94} . \log P_{Ti} + a_{10,4} . \log p_{Ni} + a_{11,4} .\log p_{Ki} + u_{i\,4} \qquad (3.3)$$

where $u_{i\,4}$ is the disturbance term for school i in this inter-relationship, and where we expect $a_{84} > 0$ and $a_{10,4} < 0$, so that the school's demand for non-teaching staff per pupil is an increasing function of its income per pupil and a decreasing function of the price that it must pay for non-teaching staff.

We may re-write (3.3) in the form:

$$b_{04} + b_{44} . \log N_i + b_{84} .\log \div_i + b_{94} . \log p_{Ti} + b_{10,4} . \log p_{Ni} + b_{11,4} .\log p_{Ki} = u_{i\,4} \qquad (3.3a)$$

where $b_{04} \equiv - a_{04}$, $b_{44} = 1$, $b_{84} \equiv - a_{84} < 0$, $b_{94} \equiv - a_{94}$, $b_{10,4} \equiv - a_{10,4} > 0$,$b_{11,4} \equiv - a_{11,4}$.

Similarly, the demand of school i for non-staff inputs per pupil, K_i, is likely to depend upon its available income per pupil, $\div_i$, upon the local price, p_{Ki} , of non-staff inputs, such as premises, and on the prices which it must pay for teaching staff and non-teaching staff expenditure out of its overall expenditure per pupil budget. In log-linear form, this implies an inter-relationship between K_i and the other variables of the form:

$$\log K_i = a_{05} + a_{85} .\log \div_i + a_{95} . \log p_{Ti} + a_{10,5} . \log p_{Ni} + a_{11,5} .\log p_{Ki} + u_{i\,5} \qquad (3.4)$$

where $u_{i\,5}$ is the disturbance term for school i in this inter-relationship. We expect $a_{85} > 0$ and $a_{11,5} < 0$, so that the school's demand for non-staff inputs per pupil is an increasing function of its income per pupil, and a decreasing function of their price.

Equation (3.4) may be re-written in the form:

$$b_{05} + b_{55}. \log K_i + b_{85} .\log \div_i + b_{95} . \log p_{Ti} + b_{10,5} . \log p_{Ni} + b_{11,5} .\log p_{Ki} = u_{i\,5} \qquad (3.4a)$$

where $b_{05} \equiv - a_{05}$, $b_{55} = 1$, $b_{85} \equiv - a_{85} < 0$, $b_{95} \equiv - a_{95}$, $b_{10,5} \equiv - a_{10,5}$,$b_{11,5} \equiv - a_{11,5} > 0$.

4. THE DEMAND FOR SCHOOL PLACES

If there are significant economies of scale in the educational production function (2.1) that links school inputs with school output, there will be a positive contribution indicated by the logarithmic coefficient $\hat{a}_6$ in (2.1) in the educational production function on the school's total size in terms of pupil numbers in contributing towards the school's educational output, q_i . An important further source of ***endogeneity*** then arises if the school's pupil numbers, m_i , themselves depend in part upon the school's level of educational attainment, q_i . The publication of school league tables and OFSTED reports, and competition between schools for pupil numbers, will place schools with more favourable examination results in a stronger position to attract increased pupil numbers. Conversely, schools with low examination results are in a weaker competitive position to increase or maintain their pupil numbers.

Such a ***demand-side*** inter-relationship arises from the desire by parents to send their children to more successful schools. The funding formulae which Local Education Authorities (LEAs) use in England to allocate resources to individual schools under the Local Management of Schools (LMS) initiative and its *Fair Funding* successor have been required to have a substantial element which depends upon the school's (age-weighted) ***pupil numbers***. This itself provides a strong incentive for each schools to accept more pupil numbers when it has the physical capacity to do so. The resulting link on the demand side between pupil numbers, m_i , and q_i , and hence log å_i , in (2.3) will again tend to undermine the assumption for the valid use of the OLS estimation technique, of zero correlation between the explanatory variables in (2.3) and the disturbance term logå_i.

The pupil numbers for school i will also depend upon its capacity, though the latter may itself change over time in response to demand pressures from local parents. It may also depend upon the local population of children of the relevant age group in the catchment area of the school. This variable may also be to some extent endogenous, reflecting in part the demand by parents for education from the school.

The demand from parents to send their children to school i may depend also upon the quality of teachers, Q_i , of the school, and its level of resourcing, as reflected in its teacher-pupil ratio, T_i , its level of non-teaching staff per pupil, N_i , its non-staff inputs per pupil, K_i , and on the general characteristics of its pupil intake, S_i . A further variable which may limit the ability of some parents to send their children to schools which are most favourably endowed with these factors is that of local house prices, p_{Hi} . If these are high, the economic ability of some parents to move into the catchment area of school i is reduced, exerting some degree of downward pressure on pupil numbers.

The inter-relationship on the demand side between pupil numbers, m_i , for school i and these other variables may then be of the form:

$$\log m_i = a_{06} + a_{16} . \log q_i + a_{26} .\log T_i + a_{36} .\log Q_i + a_{46} . \log N_i + a_{56} .\log K_i$$

$$+ a_{76} . \log S_i + a_{12,6} .\log p_{Hi} + a_{13,6} . \log \text{ø}_i + u_{i\,6} \qquad (4.1)$$

where $u_{i\,6}$ is the disturbance term for school i in this inter-relationship. We expect that $a_{16} > 0$, so that schools with superior examination performance tend to attract more pupils. Similarly we expect that $a_{26} > 0$, $a_{36} > 0$, $a_{46} > 0$, $a_{56} > 0$, $a_{76} > 0$, so that schools with more resources per pupil, higher quality teachers, and a more advantaged pupil intake tend to attract more pupils. We would also expect that

$a_{12,6} < 0$, so that higher local house prices act as an economic deterrent to locate in the local school catchment area, and that $a_{13,6} > 0$, so that the school's pupil numbers are an increasing function of the local child population density.

Equation (4.1) may be re-written in the form:

$$b_{06} + b_{16} . \log q_i + b_{26} .\log T_i + b_{36} .\log Q_i + b_{46}. \log N_i + b_{56} .\log K_i + b_{66} . \log m_i$$
$$+ b_{76} . \log S_i + b_{12,6} .\log p_{Hi} + b_{13,6} . \log ø_i = u_{i6} \qquad (4.1a)$$

where $b_{06} \equiv - a_{06}$, $b_{16} \equiv - a_{16} < 0$, $b_{26} \equiv - a_{26} < 0$, $b_{36} \equiv - a_{36} < 0$, $b_{46} \equiv - a_{46} < 0$, $b_{56} \equiv - a_{56} < 0$, $b_{66} = 1$, $b_{76} \equiv - a_{76} < 0$, $b_{12,6} \equiv - a_{12,6} > 0$,$b_{13,6} \equiv - a_{13,6} < 0$.

There may also be demand-side influences which establish a relationship between the background characteristics of the pupil intake, S_i , which a school attracts and its level of examination results, q_i , and its level of resources. Middle-class parents may be more conscious of the examination results of different schools in published school performance tables and OFSTED reports, and place greater importance on a school's examination performance and level of resources, than parents in less favourable socio-economic circumstances. They will also tend to have a greater economic ability to compete in the ***housing market*** to locate in the catchment areas of the schools with superior levels of examination performance and resourcing. Schools in a strong relative position in their examination results and their levels of resources may also tend to discriminate in favour of admitting pupils from more advantaged backgrounds in their pursuit of higher positions in published league tables of examination results.

The local nature of the public good which education typically provides implies here that demand side relationships resulting from the mobility of consumers between different local public good providers must be taken into account alongside the supply-side educational production function in the determination of equilibrium outcomes. Whilst there may not be a simple direct fiscal mechanism at work which matches local willingness to pay for the local public good to local revenue to fund provision of the public good, as in Tiebout (1956), there still nevertheless exist in the UK the intermediary mechanisms of local house prices, and school funding formulae which make school revenue dependent in large part upon pupil demand. These mechanisms will themselves ensure that demand-side factors, related in part to a school's existing level of examination performance, are likely to have an important influence upon equilibrium outcomes in the production of the local public good of education.

Links from examination results and resourcing levels to the socio-economic background of the pupil intake will establish a further source of ***endogeneity*** in the variables, here the characteristics of the pupil background, which enter the educational production function (2.1). Again the result is likely to be a breach of the key assumption of zero correlation between these variables and the level of q_i, and hence of the disturbance term $\log å_i$ in (2.3).

The dependence of the background characteristics of the pupil intake, S_i, which school i attracts upon these other factors may take on the form:

$$\log S_i = a_{07} + a_{17} . \log q_i + a_{27} .\log T_i + a_{37} .\log Q_i + a_{47} . \log N_i + a_{57} .\log K_i$$

$$+ a_{67} . \log m_i + a_{12,7} .\log p_{Hi} + u_{i\,7} \qquad (4.2)$$

where $u_{i\,7}$ is the disturbance term for school i in this inter-relationship. We expect that $a_{17} > 0$, $a_{27} > 0$, $a_{37} > 0$, $a_{47} > 0$, $a_{57} > 0$, $a_{67} > 0$, so that superior examination performance by a school, more favourable school resourcing, and the more specialised facilities which a larger school can offer, all tend to attract more advantaged pupils. We also expect that $a_{12,7} > 0$, so that higher local house prices tend to discourage parents with lower incomes from locating in the local catchment area of the school, and tend to encourage a more advantaged pupil intake into the school.

Equation (4.2) may be re-written in the form:

$$b_{07} + b_{17} . \log q_i + b_{27} .\log T_i + b_{37} .\log Q_i + b_{47} . \log N_i + b_{57} .\log K_i + b_{67} . \log m_i$$

$$+ b_{77} . \log S_i + b_{12,7} .\log p_{Hi} = u_{i\,7} \qquad (4.2a)$$

where $b_{07} \equiv - a_{07}$, $b_{17} \equiv - a_{17} < 0$, $b_{37} \equiv - a_{37} < 0$, $b_{47} \equiv - a_{47} < 0$, $b_{57} \equiv - a_{57} < 0$, $b_{67} \equiv - a_{67} < 0$, $b_{77} = 1$, $b_{12,7} \equiv - a_{12,7} < 0$.

5. THE DETERMINATION OF SCHOOL INCOME

A further important inter-relationship between the variables arises through the process which determines each school's income. This resource allocation process involves firstly the ***Standard Spending Assessment*** (SSA) allocation from central government to local authorities, and secondly what is now the ***Fair Funding*** formulae for allocating resources from Local Education Authorities (LEAs) to individual schools (see Mayston and Jesson, 1999). For earlier years, it was the Local Management of Schools (LMS) formulae that were used to allocate resources from Local Education Authorities (LEAs) to individual schools. For schools that were Grant Maintained (GM) in these earlier years, the Funding Agency for Schools (FAS) that allocated funds directly to GM schools used a funding formulae that paralleled that of the local LEA in whose area the GM school was located. Each school's income will also include the receipt of resources via specific grants, such as through the Standards Fund.

We will assume here that the income per pupil, $\div_i$, of each individual school i is determined to be at some constant level plus adjustments that depends upon the characteristics of its pupil intake, S_i , upon the size, m_i , of its pupil intake, upon the education achievement, q_i , of the school and upon the local prices, p_{Ti} , p_{Ni} and p_{Ki} which school i faces for its resource inputs. For the sake of concreteness, we will assume that the income per pupil for school i is determined by the relationship:

$$\log \div_i = a_{08} + a_{18}.\log q_i + a_{68} . \log m_i + a_{78} . \log S_i + a_{98} . \log p_{Ti} + a_{10,8} . \log p_{Ni} + a_{11,8} .\log p_{Ki} + u_{i\,8} \qquad (5.1)$$

where $u_{i\,8}$ is a random disturbance term for school i. If schools with smaller pupil intakes, with more disadvantaged pupil intakes, with relatively low examination performances, or facing higher local prices, receive more favourable resourcing for their expenditure per pupil, either through the resource allocation formulae or through the allocation of specific grants, or both, we would expect a_{18}, a_{68} and a_{78} to be negative, and a_{98}, $a_{10,8}$ and $a_{11,8}$ to be positive.

Equation (5.1) may be written in the form:

$$b_{08} + b_{18}.\log q_i + b_{68} . \log m_i + b_{78} . \log S_i + b_{88} . \log \div_i + b_{98} . \log p_{Ti} + b_{10,8} . \log p_{Ni} + b_{11,8} .\log p_{Ki} = u_{i\,8} \qquad (5.1a)$$

where $b_{08} \equiv - a_{08}$, $b_{18} \equiv - a_{18} > 0$, $b_{68} \equiv - a_{68} > 0$, $b_{78} \equiv - a_{78} > 0$, $b_{88} = 1$, $b_{98} \equiv -a_{98} < 0$, $b_{10,8} \equiv - a_{10,8} < 0$, $b_{11,8} \equiv - a_{11,8} < 0$.

6. THE DETERMINATION OF LOCAL PRICES

There may also be further links between the above variables because of a dependency of the local salaries which are required to attract teaching staff and non-teaching staff to school i upon local house prices. Similarly premises costs that enter into non-staff costs, p_{Ki} , may be dependent to some extent upon local property prices. This may imply relationships of the form:

$$\log p_{Ti} = a_{09} + a_{12,9} .\log p_{Hi} + u_{i\,9} \tag{6.1}$$

$$\log p_{Ni} = a_{0,10} + a_{12,10} .\log p_{Hi} + a_{14,10} . \log Ø_i + u_{i,\,10} \tag{6.2}$$

$$\log p_{Ki} = a_{0,11} + a_{12,11} .\log p_{Hi} + u_{i,\,11} \tag{6.3}$$

where $u_{i\,9}$, $u_{i,\,10}$, and $u_{i\,,11}$ are the relevant disturbance terms for school i, and where we expect $a_{12,9} > 0$, $a_{12,10} > 0$, and $a_{12,11} > 0$. $Ø_i$ is here the local unemployment rate, which may have a negative influence on non-teaching staff costs, with $a_{14,10} < 0$.

Equations (6.1) - (6.3) may be re-written in the form:

$$b_{09} + \log p_{Ti} + b_{12,9} .\log p_{Hi} = u_{i\,9} \tag{6.1a}$$

$$b_{0,10} + \log p_{Ni} + b_{12,10} .\log p_{Hi} + b_{14,10} . \log Ø_i = u_{i,\,10} \tag{6.2a}$$

$$b_{0,11} + \log p_{Ki} + b_{12,11} .\log p_{Hi} = u_{i,\,11} \tag{6.3a}$$

where $b_{0,\,9} \equiv - a_{0,\,9}$, $b_{12,9} \equiv - a_{12,\,9} < 0$, $b_{0,\,10} \equiv - a_{0,\,10}$, $b_{12,10} \equiv - a_{12,\,10} < 0$, $b_{14,10} \equiv - a_{14\,,10} > 0$, $b_{0,\,11} \equiv - a_{0,\,11}$, $b_{12,11} \equiv - a_{12,\,11} < 0$.

Finally, local house prices, p_{Hi} , may be higher than otherwise for schools with higher levels of resources and superior examination results, because of demand pressures from parents seeking to move into the catchment area of such schools pushing up house prices. They may also be influenced by the local socio-economic background characteristics. High levels of unemployment, housing density, crime and other social problems may tend to depress local house prices. House prices may also be affected by other local factors, such as the distance of the school from central London, d_i , and the distance of the school, $ë_i$, from the nearest city other than London (normalising on $ë_i = 1$ for cases where the nearest city is London). These distance variables will be assumed here not to form part of the local pupil background characteristics which are relevant to the educational production function (2.1).

The overall influences on local house prices, p_{Hi} , may then be of the form:

$$\log p_{Hi} = a_{0,12} + a_{1,12} . \log q_i + a_{2,12} .\log T_i + a_{3,12} .\log Q_i + a_{4,12}. \log N_i + a_{5,12} .\log K_i + a_{6,12}.\log m_i + a_{7,12} . \log S_i + a_{13,12}. \log Ø_i + a_{15,12}. \log d_i + a_{16,12}. \log ë_i + u_{i,12} \tag{6.4}$$

where $u_{i\,12}$ is the disturbance term for school i in this relationship, and we expect $a_{1,12} > 0$, $a_{2,12} > 0$, $a_{3,12} > 0$, $a_{4,12} > 0$, $a_{5,12} > 0$, $a_{6,12} > 0$, $a_{7,12} > 0$, $a_{13,12} < 0$ and $a_{15,12} < 0$, $a_{16,12} < 0$.

Equation (6.4) may be re-written in the form:

$$b_{0,12} + b_{1,12} . \log q_i + b_{2,12} .\log T_i + b_{3,12} .\log Q_i + b_{4,12}. \log N_i + b_{5,12} .\log K_i + b_{6,12}.\log m_i + b_{7,12} . \log S_i + b_{12,12} .\log p_{Hi} + b_{13,12}. \log Ø_i + b_{15,12}. \log d_i + b_{16,12}. \log ë_i = u_{i,12} \tag{6.4a}$$

where $b_{0,\,12} \equiv - a_{0,\,12}$, $b_{1,12} \equiv - a_{1,12} < 0$, $b_{2,\,12} \equiv - a_{2,\,12} < 0$, $b_{3,12} \equiv - a_{3,12,} < 0$, $b_{4,12} \equiv -a_{4,12,} < 0$, $b_{5,12} \equiv - a_{5,12,} < 0$, $b_{6,12} \equiv - a_{6,12,} < 0$, $b_{7,12} \equiv - a_{7,12,} < 0$, $b_{12,12} = 1$, $b_{13,12} \equiv - a_{13\,,12} > 0$, $b_{15,12} \equiv - a_{15\,,12} > 0$, $b_{16,12} \equiv - a_{16\,,12} > 0$.

7. THE EXTENT OF THE ENDOGENEITY BIAS

The equations (2.3a) - (6.4a) form a ***simultaneous equations model***, here with twelve structural equations to the model involving twelve endogenous variables and four exogenous variables. The twelve endogenous variables are the logarithms of the following variables:

1. q_i = school i’s educational output, as reflected in the examination performance of its pupils
2. T_i = school i’s teacher-pupil ratio
3. Q_i = a measure of teacher quality for school i
4. N_i = non-teaching staff per pupil in school i

5. K_i = non-staff expenditure in volume terms in school i
6. m_i = total pupil numbers for school i
7. S_i = a measure of the socio-economic background of the pupil intake into school i
8. $÷_i$ = school i’s income per pupil
9. p_{Ti} = the local pay level for teachers of a standard quality facing school i
10. p_{Ni} = the local pay level for non-teaching staff facing school i
11. p_{Ki} = the local level of non-staff costs facing school i
12. p_{Hi} = the local level of house prices in the catchment area of school i.

The four exogenous variables are the logarithms of the following variables:

13. $ø_i$ = the local population density of children in the relevant age groups
14. $Ø_i$ = the local unemployment rate in the catchment area of school i
15. d_i = the distance of school i from central London
16. $ë_i$ = the distance of the school from the nearest city other than London (normalising on $ë_i$ = 1 for cases where the nearest city is London).

The resulting simultaneous equations model is of the general form:

$$\sum_{k=1}^{n'} y_{ik}.b_{kj} + \sum_{h=1}^{n''} z_{ih}.ô_{hj} = u_{ij} \qquad for\ \ i=1,...,v;\ j=1,...,n' \tag{7.1}$$

where y_{ik} denotes the ith observation on the kth endogenous variable for each of the n’ endogenous variables, and z_{ih} denotes the ith observation on the hth pre-determined variable for each of the n” pre-determined variables, which may include exogenous and lagged endogenous
variables. The b_{kj} and $ô_{hj}$ are the corresponding structural parameters in the jth structural relation of the above simultaneous equations model. u_{ij} is the random disturbance term for the ith observation in the jth structural relation in the model (7.1) for each of v observations. In matrix form, (7.1) may be written as:

$$Y.B + Z.Ã = U \tag{7.1a}$$

where $Y \equiv [y_{ik}]$, $B \equiv [b_{kj}]$, $Z \equiv [z_{ih}]$, $Ã \equiv [ô_{hj}]$ and $U \equiv [u_{ij}]$.

In the model (2.3a) - (6.4a), we have n’ = 12 endogenous variables, with the y_{ik} given by the logarithms of the variables 1 - 12 listed above. We also have n” = 4 exogenous variables, with the $z_{i\,h}$ given by the logarithms of the variables 13,..., 16 listed above.

We will assume that the disturbance terms u_{ij} in the structural equations are independently and identically multi-variate normally distributed for each observation i=1,....,v, with zero means and a diagonal covariance matrix V. The covariance matrix is assumed to have positive diagonal elements $ó_j^2$ for j=1,...,n’ and zero off-diagonal terms, reflecting uncorrelated disturbances across the different structural relations within the model. The disturbance terms in each individual structural equation within the complete model are also assumed to be uncorrelated with the values of each of the observed explanatory variables in (7.1).

The bias that results from neglecting the simultaneous equations nature of the problem and applying the standard Ordinary Least Squares (OLS) multivariate regression technique to directly to estimating the parameters $â_2$, ..., $â_7$ of the educational production function equation (1.3) is shown in the Appendix to be of the form:

$$\text{plim } è_{1k} \equiv \text{plim}\hat{â}_k - â_k \quad \text{for } k = 2,\ldots, n' \tag{7.2}$$

$$= -\frac{ó_1^2}{î} \sum_{j=2}^{n'} b_{1j}\,(â_k.b_{1j} + b_{kj}) \;/\; ó_j^2 \tag{7.3}$$

$$\textit{where} \quad î \equiv \left(1 + \sum_{j=2}^{n'} b_{1j}^2 . \frac{ó_1^2}{ó_j^2}\right) > 0 \tag{7.4}$$

$\text{plim}\hat{â}_k$ here denotes the estimated coefficient which the application of OLS will achieve for the coefficient $â_k$ as the number of observations increases to infinity. This means that it excludes biases which are due simply to a small sample size. As in equation (2.1) above, each $â_k$ denotes the (logarithmic) contribution of the each relevant school input variable to the school’s educational output, q_i . We expect the $â_k$ to be positive in value. From our above discussion, we expect that:

$$b_{12} = 0,\ b_{13} < 0,\ b_{14} = 0,\ b_{15} = 0,\ b_{16} < 0,\ b_{17} < 0,\ b_{18} > 0,\ b_{19} = 0,\ b_{1,10} = 0,\ b_{1,11} = 0,\ b_{1,12} < 0 \tag{7.5}$$

$$b_{22} = 1,\ b_{23} < 0,\ b_{24} = 0,\ b_{25} = 0,\ b_{26} < 0,\ b_{27} = 0\ ,\ b_{28} = 0,\ b_{29} = 0,\ b_{2,10} = 0\ ,\ b_{2,11} = 0,\ b_{2,12} < 0 \tag{7.6}$$

The signs of the coefficients in (7.5) and (7.6), together with (7.3) and (7.4), imply that

$$\text{plim } \grave{e}_{12} \equiv \text{plim}\hat{\hat{a}}_2 - \hat{a}_2 < 0 \qquad (7.7)$$

so that the OLS estimator, $\hat{\hat{a}}_{2,}$ ***under-estimates*** the true value of $\hat{a}_2$, the proportionate contribution which the teacher-pupil ratio makes to school output. From (7.3), the extent of the under-estimate is, moreover, ***the sum*** of the endogeneity biases which arise from all those equations in the model for which the relevant b_{1j} coefficients on the school's educational performance, q_i , in (7.5) are non-zero.

These additive ***sources of endogeneity*** include not only the eighth equation in (5.1) that links school funding through the resource allocation process to the school's level of examination success. They also include the third equation that links the ability of a school to attract teachers of higher quality to the level of its examination performance, the sixth equation that links the demand for school places by parents and pupils to the school's examination performance, the seventh equation that links the socio-economic background of the pupils that the school attracts to the level of the school's examination performance, and the twelfth equation that links local house prices to the level of the school's examination performance.

All of these additional inter-relationships will add to the extent of the downward ***endogeneity bias*** which result in OLS under-estimating the logarithmic coefficient, $\hat{a}_2$, of the proportionate contribution which the teacher-pupil teacher ratio makes to school output. The sixth, seventh and twelfth equations, moreover, reflect ***demand-side*** factors from parents boosting demand for those schools with high levels of school educational output, in addition to the demand-side factors which arise from the school funding mechanisms. As noted below, it is possible that these latter sources of endogeneity bias in estimating the contributions of school resource variables to school output will be at least as great as that from the funding mechanisms.

We also have:

$$b_{42} = 0,\ b_{43} < 0,\ b_{44} = 1,\ b_{45} = 0,\ b_{46} < 0,\ b_{47} < 0\ ,\ b_{48} =0,\ b_{49} = 0,\ b_{4,10} = 0\ ,\ b_{4,11} = 0,\ b_{4,12} < 0 \qquad (7.8)$$

$$b_{52} = 0,\ b_{53} < 0,\ b_{54} = 0,\ b_{55} = 1,\ b_{56} < 0,\ b_{57} < 0,\ b_{58} =0,\ b_{59} = 0,\ b_{5,10} = 0\ ,\ b_{5,11} = 0,\ b_{5,12} < 0 \qquad (7.9)$$

which together with (7.2) - (7.4) can be shown to imply that:

$$\text{plim } \grave{e}_{14} \equiv \text{plim}\hat{\hat{a}}_4 - \hat{a}_4 < 0 \quad \text{and} \quad \text{plim } \grave{e}_{15} \equiv \text{plim}\hat{\hat{a}}_5 - \hat{a}_5 < 0 \qquad (7.10)$$

i.e. ***downward bias*** in the OLS estimate of the contribution of the other two resource inputs to the educational production function of non-teaching staff per pupil, N_i , and non-staff inputs per pupil, $K_{i.}$

In the case of the quality of teachers variable, Q_i , we have:

$$b_{32} > 0,\ b_{33} = 1,\ b_{34} = 0,\ b_{35} = 0,\ b_{36} < 0,\ b_{37} < 0\ ,\ b_{38} =0,\ b_{39} = 0,\ b_{3,10} = 0\ ,\ b_{3,11} = 0,\ b_{3,12} < 0 \qquad (7.11)$$

Again downward bias in the OLS estimate of the contribution of the quality of teachers variable, Q_i , in the educational production function is indicated by all terms in (7.3), given (7.4) , (7.5) and (7.11), except for the term involving the third equation j = 3. Here we have $b_{13} < 0$ and $b_{33} = 1$, making this term of uncertain sign for $\hat{a}_3 > 0$. Since in equations (2.1) and (3.2a) respectively:

$$\hat{a}_3 = \partial \log q_i / \partial \log Q_i \quad \text{and} - b_{13} = \partial \log Q_i / \partial \log q_i \qquad (7.12)$$

stability requires that:

$$- \hat{a}_3 .b_{13} < 1 \qquad (7.13)$$

If condition (7.13) does not hold, there will be a ***positive feedback effect*** of schools with higher examination results attracting better quality teachers which in turn contribute towards higher examination results in the educational production function. This will tend to lead to an ***unstable cumulative advantage*** for schools with initially higher examination results. If (7.13) holds, the term involving the third equation j = 3 in (7.3) will offset to some extent the downward bias in the OLS estimate of $\hat{a}_3$ that otherwise prevails. If (7.13) does not hold, and a process of unstable cumulative advantage does potentially exist, then the term involving the third equation j = 3 in (7.3) will reinforce the overall downward bias in the OLS estimate of $\hat{a}_3$ that the other terms of (7.3) imply.

For the remaining variables of pupil numbers, m_i , and socio-economic background, S_i , that also enter the educational production function, we have:

$$b_{62} = 0,\ b_{63} < 0,\ b_{64} = 0,\ b_{65} = 0,\ b_{66} = 1,\ b_{67} < 0,\ b_{68} > 0,\ b_{69} = 0,\ b_{6,10} = 0,\ b_{6,11} = 0,\ b_{6,12} < 0 \qquad (7.14)$$

$$b_{72} = 0,\ b_{73} < 0,\ b_{74} = 0,\ b_{75} = 0,\ b_{76} < 0,\ b_{77} =1,\ b_{78} > 0,\ b_{79} = 0,\ b_{7,10} = 0,\ b_{7,11} = 0,\ b_{7,12} < 0 \qquad (7.15)$$

so that similar remarks apply as for teacher quality. There will in general be a ***downward bias*** in the estimates which OLS achieves for the contribution which pupil numbers, m_i , and socio-economic background, S_i , make to the educational production function. However, when the stability conditions

$$- \hat{a}_6 .b_{16} < 1 \qquad \text{and} \qquad - \hat{a}_7 .b_{17} < 1 \qquad (7.16)$$

hold, this downward bias will be reduced to some extent by some degree of positive offset from the terms in (7.3) corresponding to j = 6 and j = 7 respectively.

If (7.16) does not hold, an unstable cumulative advantage potentially exists from schools with better examination results attracting greater demand from parents, in terms of pupil numbers and pupils with more advantaged pupil backgrounds, which in turn boost examination performance in an unstable way. In such a case, there will be no positive offset to the downward bias in the estimates which OLS achieves here. The potential instability of cumulative advantage may also be offset in practice by the funding mechanism for schools protecting to some extent the expenditure per pupil of schools with lower pupil intakes and with more disadvantaged pupil intakes.

8. CONCLUSIONS

There are several important additional possible inter-relationships between the variables which enter into the school-level educational production function which are likely to result in the standard Ordinary Least Squares estimation technique ***under-estimating*** the true influence of school resourcing variables on pupil educational outcomes. The inter-relationships include ***demand-side*** relationships that make the demand for school places by parents and pupils, the socio-economic background of the pupils that the school attracts, and local house prices, themselves functions of the school's degree of examination success. In addition, they include the inter-relationship between the ability of a school to attract teachers of higher quality in the labour market and the school's standing in school league tables of examination performance. They also include possible links between the level of school funding through the resource allocation process and the existing level of the school's examination success.

Under reasonable assumptions, the effect of these additional inter-relationships will be to ***bias downwards*** the estimates which OLS makes of the influence of school resource variables on pupils' educational outcomes in a ***cumulative*** way. Little faith can then be placed upon conclusions based upon the OLS estimations contained in many existing studies of the educational production function that the underlying influence of school resource variables on pupil educational outcomes are not significantly positive.

While we have analysed the associated endogeneity problem in terms of the OLS estimation of what will be a 'statistically average' educational production function, similar conclusions are likely to hold for the estimation of ***stochastic frontier*** production function models (see Aigner *et al*, 1997; Forsund *et al*,

1980) in the presence of the additional inter-relationships between the underlying variables which we have analysed above. Similarly, the use of ***multilevel***, or ***error component***, models (see Goldstein, 1987) that would recognise the contributions of variables at different levels of the education process, such as the LEA, school and pupil levels, represents a variation on the basic single-equation multivariate regression model that is unlikely to escape the biases that are introduced by the additional inter-relationships we have discussed above. The estimation problems which arise in the context of ***non-parametric*** models of the educational production frontier, such as those of Data Envelopment Analysis (DEA), are discussed in Mayston (2000b), and are compounded by the existence of additional inter-relationships between the variables beyond that described by the underlying production frontier.

The policy importance of correctly estimating the educational production function underlines the need for improvements in the database of comparative information between schools, including resource variables, that are advocated in Mayston and Jesson (1999). The scope for making progress in the empirical estimation of the educational production function based upon our above model of the inter-relationships that are likely to exist between the key variables will be discussed in detail in a later paper.

APPENDIX

Our system of structural equations is of the form:

$$\sum_{k=1}^{n^{/}} y_{ik}.b_{kj} \ + \ \sum_{h=1}^{n^{//}} z_{ih}.\hat{o}_{hj} \ = \ u_{ij} \qquad \mathit{for}\ \ i=1,...,v;\ j=1,...,n^{/} \tag{A.1}$$

where y_{ik} denotes the ith observation on the kth endogenous variable for each of the n’ endogenous variables, and z_{ig} denotes the ith observation on the hth predetermined variable for each of the n” predetermined variables, which may include exogenous and lagged endogenous variables. The b_{kj} and $\hat{o}_{hj}$ are the corresponding structural parameters in the jth structural relation of the above structural model. u_{ij} is the random disturbance term for the ith observation in the jth structural relation in the model (A.1). The u_{ij} are assumed to be independently and identically multi-variate normally distributed for each observation i=1,....,v, with zero means and a diagonal covariance matrix V, with positive diagonal elements $\acute{o}_j^2$ for j=1,...,n’ and uncorrelated disturbances across different structural relations within the model. The disturbance terms are also assumed to be uncorrelated with the values of each of the observed explanatory variables in (A.1).

The model (A.1) may be written in matrix form as:

$$YB + Z\tilde{A} = U \tag{A.2}$$

where Y is the v x n’ matrix with elements y_{ik} , B is the n’ x n’ matrix with elements b_{kj} , Z is the v x n” matrix with elements $z_{i\,h}$, $\tilde{A}$ is the n” x n’ matrix with elements $\hat{o}_{hj}$. We will assume that B is a non-singular matrix. We may write:

$$Y = [\ y \ \ Y_0] \text{ where } y \equiv [y_{i1}] \text{ and } Y_0 \equiv [y_{ik}] \text{ for } i=1,..., v \text{ and } k = 2,...,n' \tag{A.3}$$

$$X \equiv [Y \ \ Z] \equiv [x_{ik}] \text{ for } i=1,..., v \text{ and } k = 1,...,n \ \text{ where } n \equiv n' + n'' \tag{A.4}$$

$$X_0 \equiv [Y_0 \ \ Z] \quad \text{and } A' \equiv [B' \quad \tilde{A}'] \tag{A.5}$$

(A.2) may then be written:

$$XA = U \tag{A.6}$$

If we normalise the structural equations by setting $b_{11} = 1$, we may write:

$$B = \begin{bmatrix} 1 & a \\ c & B_0 \end{bmatrix} \quad \text{where } B_0 \equiv [\, b_{kj} \,] \text{ for } k = 2,...,n' \text{ ; } j = 2,...,n' \tag{A.7}$$

$$a \equiv [b_{1j}\ ] \text{ for } j = 2,...,n' \tag{A.8}$$

$$c \equiv [b_{k1}\ ] = - [\hat{a}_k] \text{ for } k = 2,...,n' \tag{A.9}$$

when we set:

$$\hat{a}_k \equiv - b_{k1} \quad \text{for } k = 2,...,n' \tag{A.10}$$

Setting:

$$\hat{a}_{n' + h} \equiv - \hat{o}_{h1} \quad \text{for } h = 1,....,n'' \tag{A.11}$$

yields from (A.2) - (A.11), the first structural relation of the equation written in the form:

$$y = X_0\, \hat{a} + u \tag{A.12}$$

where $\hat{a} \equiv [\ \hat{a}_k\ ]$ for k = 2, ..., n and $u \equiv [\ u_{i1}\ ]$ for i =1,..., v.

We have the OLS estimator of the coefficients of (A.12) given by:

$$\hat{\hat{a}} = (X_0' X_0)^{-1} X_0' y \tag{A.13}$$

as in Johnston (1984, p. 171). From (A.12):

$$\hat{\hat{a}} = (X_0' X_0)^{-1} X_0' X_0 \hat{a} + (X_0' X_0)^{-1} X_0' u \tag{A.14}$$

$$= \hat{a} + (X_0' X_0)^{-1} X_0' u \tag{A.15}$$

The extent of the bias which the use of Ordinary Least Squares (OLS) produces in its estimates of the coefficients of the first structural relation (A.12), away from their true underlying values given by the vector $\hat{a}$, is then given by:

$$\grave{e} \equiv \hat{\hat{a}} - \hat{a} = (X_0' X_0)^{-1} X_0' u \tag{A.16}$$

From (A.5), we have:

$$X_0' X_0 = \begin{bmatrix} Y_0' Y_0 & Y_0' Z \\ Z' Y_0 & Z'Z \end{bmatrix} \tag{A.17}$$

Let

$$(X_0' X_0)^{-1} \equiv \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \tag{A.18}$$

where P is (n’ -1) x (n’ -1), Q is (n’ - 1) x n”, R is n” x (n’ - 1), and S is n” x n”. From Hadley (1961, p. 109), we have:

$$P = H^{-1} \text{ where } H \equiv Y_0'Y_0 - Y_0' Z (Z'Z)^{-1}Z'Y_0 \tag{A.19}$$

$$Q = - PY_0' Z (Z'Z)^{-1} \tag{A.20}$$

$$R = - (Z'Z)^{-1} Z'Y_0 P \tag{A.21}$$

$$S = (Z'Z)^{-1} - (Z'Z)^{-1} Z'Y_0 Q \tag{A.22}$$

From (A.2):

$$Y = UB^{-1} - Z\tilde{A} B^{-1} \tag{A.23}$$

$$\text{Let} \quad B^{-1} \equiv \begin{bmatrix} ç & g \\ ö & F \end{bmatrix} \tag{A.24}$$

where ç is 1 x 1, g is 1 x (n’ - 1), ö is (n’ - 1) x 1 , and F is (n’ - 1) x (n’ - 1). From (A.7) and Hadley (1961, p. 109), we have:

$$F = (B_0 - ca)^{-1} \quad , ö = - Fc , \quad g = -aF , \quad ç = 1 - aö \tag{A.25}$$

$$\text{Let} \quad G' \equiv [\ g' \quad F'\ ] = F' [-a' \quad I] \tag{A.26}$$

Then from (A.3), (A.23) - (A.24):

$$Y_0 = UG - Z\tilde{A}G \tag{A.27}$$

$$Y_0' Z = G'U'Z - G'\tilde{A}'Z'Z \tag{A.28}$$

$$Y_0' Z (Z'Z)^{-1} = G'U'Z (Z'Z)^{-1} - G'\tilde{A}' \tag{A.29}$$

$$Z'Y_0 = Z'UG - Z'Z\tilde{A}G \tag{A.30}$$

$$Y_0' Z (Z'Z)^{-1} Z'Y_0 = G'U'Z (Z'Z)^{-1} Z'UG - G'U'Z\tilde{A}G$$

$$- G'\tilde{A}' Z'UG' + G'\tilde{A}'Z'Z\tilde{A}G \quad (A.31)$$

$$Y_0' Y_0 = G'U'UG - G'\tilde{A}'Z'UG - G'U'Z\tilde{A}G + G'\tilde{A}'Z'Z\tilde{A}G \quad (A.32)$$

Hence from (A.19), (A.31) - (A.32):

$$H = G'U'UG - G'U'Z (Z'Z)^{-1}Z'UG \quad (A.33)$$

Since the disturbance terms are assumed to be uncorrelated with the explanatory variables in (A.1), we have:

$$\text{plim } U'Z = 0, \text{ plim } Z'U = 0 \text{ and } \text{plim}Z'u = 0 \quad (A.34)$$

where plim denotes the value of the term as the number of observations, m, goes to infinity. Hence from Johnston (1984, pp. 269 - 271):

$$\text{plim } H = G'VG \quad \text{where } V \equiv [\ \acute{o}_j^2.\ddot{a}_{j\,k}] \text{ for } j,k = 1,..., n' \quad (A.35)$$

and where $\ddot{a}_{j\,k} = 1$ for $j = k$ and $\ddot{a}_{j\,k} = 0$ for $j \neq k$. Hence from (A.26):

$$\text{plim } H = F'(a'\acute{o}_1^2\ a + V_0)\ F \text{ where } V_0 \equiv [\ \acute{o}_j^2.\ddot{a}_{j\,k}] \text{ for } j,k = 2,..., n' \quad (A.36)$$

From (A.19), (A.36) and Johnston (1984, p. 271), we have:

$$P^* \equiv \text{plim } P = \text{plim } H^{-1} = F^{-1}\ (a'\acute{o}_1^2\ a + V_0)^{-1}\ (F')^{-1} \quad (A.37)$$

From (A.21) and (A.30), we have:

$$R = -\ (Z'Z)^{-1}\ Z'UGP + \tilde{A}GP \quad (A.38)$$

Hence from (A.34) and (A.37):

$$R^* \equiv \text{plim } R = \tilde{A}GP^* = \tilde{A}GF^{-1}\ (a'\acute{o}_1^2\ a + V_0)^{-1}\ (F')^{-1} \quad (A.39)$$

From (A.5), (A.16), (A.18) and (A.27):

$\wedge$

$$è_1 \equiv [â_k - â_k] \text{ for } k = 2,..., n' \qquad (A.40)$$

$$= PY_0' u + QZ'u \qquad (A.41)$$

$$= P (G'U'u - G'Ã'Z'u) + QZ'u \qquad (A.42)$$

$$è_2 \equiv [\hat{â}_k - â_k] \text{ for } k = n'+1,..., n \qquad (A.43)$$

$$= RY_0' u + S Z'u \qquad (A.44)$$

$$= R (G'U'u - G'Ã'Z'u) + S Z'u \qquad (A.45)$$

Since the u_{ij} are assumed to be uncorrelated across the different structural relations:

$$\text{plim } U'u = w \qquad \text{where } w' \equiv (ó_1^2 , 0, 0, ...,0) \qquad (A.46)$$

Hence from (A.25) - (A.26), (A.34), (A.37) - (A.45):

$$\text{plim } è_1 = [\text{plim}\hat{â}_k - â_k] \text{ for } k = 2,..., n' \qquad (A.47)$$

$$= P^*G'w \qquad (A.48)$$

$$= F^{-1} (a'ó_1^2 a + V_0)^{-1} (F')^{-1} G'w \qquad (A.49)$$

$$= (B_0 - ca)(a'ó_1^2 a + V_0)^{-1} (F')^{-1} F' [-a' \quad I] w \qquad (A.50)$$

$$= (ca - B_0)Ka'ó_1^2 \quad \text{where } K \equiv (a'ó_1^2 a + V_0)^{-1} \qquad (A.51)$$

We may show that:

$$K = [(ä_{jk}/ ó_k^2) - ((b_{1j} .b_{1k} .ó_1^2 / (ó_j^2 .ó_k^2 . î))] \qquad (A.52)$$

$$\textit{where} \quad î \equiv \left(1 + \sum_{j=2}^{n'} b_{1j}^2 . \frac{ó_1^2}{ó_j^2} \right) > 0 \qquad (A.53)$$

with $K (a'ó_1^2 a + V_0) = I$, using (A.8). From (A.8), (A.52) - (A.53):

$$K a' = [b_{ik} / (î . ó_k^2)] \quad \text{for } k = 2,...,n' \qquad (A.54)$$

Hence from (A.8), (A.47) and (A.51):

$$\text{plim } è_1 = [\text{plim}\hat{â}_k - â_k] \text{ for } k = 2,..., n' \qquad (A.55)$$

$$= (ca - B_0) [b_{ik} / (î . ó_k^2)] ó_1^2 \qquad (A.56)$$

$$= (-\hat{a}_k \,.\, b_{1j} \; - b_{kj})[\, b_{ik} / (\hat{\imath} \,.\, \acute{o}_k^2\,)]\acute{o}_1^2 \tag{A.57}$$

Hence from (A.57) the extent of the asymptotic bias between the OLS estimator $\hat{\hat{a}}_k$ for the kth endogenous variable in the structural equation (A.12) and its true underlying value $\hat{a}_k$ that remains when the number of observations increases to infinity is given by:

$$\text{plim}\, \grave{e}_{1k} \equiv \text{plim}\hat{\hat{a}}_k - \hat{a}_k \;\text{ for k =2,..., n'} \tag{A.58}$$

$$= -\frac{\acute{o}_1^2}{\hat{\imath}} \sum_{j=2}^{n'} b_{1j}\; (\hat{a}_k.b_{1j} \;+\; b_{kj}) \;/\; \acute{o}_j^2 \tag{A.59}$$

From (A.26), (A.34), (A.39), (A.43) - (A.46), and (A.51):

$$\text{plim}\grave{e}_2 \equiv [\text{plim}\; \hat{\hat{a}}_k - \hat{a}_k] \;\text{ for k = n'+1,..., n} \tag{A.60}$$

$$= R^*G'w \tag{A.61}$$

$$= \tilde{A}\, G\, F^{-1}\, K\, (F')^{-1} F'\; [-a' \quad I] \tag{A.62}$$

$$= \tilde{A}\; [-a' \quad I]'\; K\; a'\acute{o}_1^2 \tag{A.63}$$

$$= [\hat{o}_{h1}\, b_{1j} \; - \hat{o}_{hj}\,]\; [\, b_{ik} / (\hat{\imath} \,.\, \acute{o}_k^2\,)]\; \acute{o}_1^2 \tag{A.64}$$

using (A.2), (A.8) and (A.54).

Hence from (A.64) the extent of the asymptotic bias between the OLS estimator $\hat{a}_k$ for the kth predetermined variable in the structural equation (A.12) and its true underlying value $\hat{a}_k$ that remains when the number of observations increases to infinity is given by:

$$\text{plim}\grave{e}_2 \equiv \text{plim}\; \hat{\hat{a}}_k - \hat{a}_k \;\text{ for k = n'+1,..., n} \tag{A.65}$$

$$= -\frac{\acute{o}_1^2}{\hat{\imath}} \left(\sum_{j=2}^{n'} \frac{b_{1j}}{\acute{o}_j^2} \;(\hat{a}_{n'+h} b_{1j} \;+\; \hat{o}_{hj}) \right) \tag{A.66}$$

using (A.11).

REFERENCES

Aigner, D., C. A. K. Lovell and P.Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, vol. 6, pp. 21 - 37.

Audit Commission (1993 - 1996), *Adding Up the Sums 1 - 4*. HMSO: London.

Department for Education and Employment (1998), *Fair Funding: Improving Delegation to Schools*. London: DfEE.

Forsund, F. R. , C.A.K. Lovell and P. Schmidt (1980), "A Survey of Frontier Production Functions and of their Relationship to Efficiency Measurement", *Journal of Econometrics*, vol. 13, pp. 5 - 25.

Goldstein, H. (1987), *Multilevel Models in Education and Social Research*. Charles Griffin and Company: London.

Gujarati, D. (1995), *Basic Econometrics*, 3rd edition. McGraw-Hill: London.

Hadley, G. (1961), *Linear Algebra.* Addison-Wesley: London.

Hanushek, E. (1979), "Conceptual and Empirical Issues in the Estimation of Educational Production Functions", *Journal of Human Resources*, vol. 14, pp. 351 - 388.

Hanushek, E. (1986), "The Economics of Schooling: Production and Efficiency in Public Schools", *Journal of Economic Literature*, vol. 24, pp. 1141 - 1177.

Johnston, J. (1984), *Econometric Methods*, 3rd edn. McGraw-Hill: London.

Mayston, D.J. (1996), "Educational Attainment and Resource Use: Mystery or Econometric Misspecification?", *Education Economics*, vol. 4, pp. 127 - 142.

Mayston, D.J. (2000a),"Accounting, Information and the Development of Evidence-based Resourcing Strategies in Education", *University of York Discussion Papers*.

Mayston, D.J. (2000b), “Performance Management and Performance Measurement in the Education Sector”,*University of York Discussion Papers*.

Mayston, D. J. and D. Jesson (1999), *Linking Educational Resourcing with Enhanced Educational Outcomes*, DfEE Research Report RR179. Department for Education and Employment: London.

Murdane, R.J. (1996), “Staffing the Nation’s Schools with Skilled Teachers”, in E.A. Hanushek and D.W. Jorgenson eds, *Improving America’s Schools: The Role of Incentives*. National Academy Press: Washington DC.

Tiebout, C. (1956), “A Pure Theory of Local Expenditure”, *Journal of Political Economy*, vol. 64, pp. 416 - 424.