A pearl of great price, the free education system of Sri Lanka
Ranasinghe, A.

Citation for published version (APA):
Ranasinghe, A. (1999). A pearl of great price, the free education system of Sri Lanka

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 7

Production, Demand and Private Cost of Education Quality:
An Application of Canonical Regression

7.1 Introduction

As noted in Chapter 2, two inter-related but distinct issues dominate the recent literature on quality of education: the education (school) production function, which measures the effects of "school inputs" on the "output" of schools, and the demand for schools. The former became popular in the United States after the controversial Coleman report in 1966. The Coleman report argued that inputs of a given school are not important determinants of the performance of pupils being trained by the school. According to this report, the most important factors are the family characteristics and the effect of peer-groups. This controversial finding has motivated many economists to further investigate the issue; they have emerged with mixed findings on the relation between inputs and output of the education system [see Hanushek (1986)].

This leaves the controversy of the Coleman report unresolved until today. One of the main problems associated with the estimation of the school production function is that the empirical estimation of the production function is based on only some observable dimension of its output; meanwhile, the true output of schools may not be measurable due to its many dimensions [Hanushek (1986)]. In general, the conceptual framework behind the school production function presumes that schools are different from each other in efficiency, whereby some schools are better than others. The immediate implication of this belief for school selection is the theme for the latter issue, demand for schools. If schools are different from each other, a rational pupil chooses the best available school. Consequently, school choice becomes an endogenous variable in the human capital model. Most literature related with the school demand function is unique in its methodology and estimation techniques. In general, authors of this literature define several school types based on the quality of service, and focus on the estimation of the willingness to pay for quality improvement, and (or) upon the welfare loss associated with fee levying education systems [see Gertler and Glewwe (1990) and Athurupana (1993) for discrete choice school demand functions and Gertler et al. (1987) Gertler et al. (1994) for applications to demand for health care]. The estimation method they have applied is the Nested-Multinomial-Logit (NMNL) of McFadden (1981), and the method for estimating willingness to pay is from Small and Rosen (1981). Nevertheless, this method is not without limitations. One can criticise that the classification of schools is highly subjective and the regression results are heavily dependent on this classification. For example, Gertler and Glewwe (1990) used the classification of faraway and nearby schools relative to the residence of the respondent.
The maintained hypothesis of this classification is that the quality of education provided by faraway schools is better than that provided by the nearby schools. Athurupana (1993) classified schools into three categories as private schools, fee levying public schools and free schools. His presumption is that private schools are better than fee levying public schools and the fee levying public schools are better than free schools. The advantage of this method is that it does not require the researcher to use an explicit school quality measure. This at least partially overcomes the problem of non-measurability of school quality. The absence of a hard and fast rule for school quality classification remains as the major drawback of this method. In fact, the final estimate is highly subject to the choice of school classification. The use of broad school categories such as the one used by Gertler and Glewwe (1990), or by Athurupana (1993) neglect the possible variation of school quality within a given category. In other words, this method assumes that schools in a given category are homogenous.

In this chapter we relax this assumption and treat each school as a separate entity. The quality of the service then differs from school to school. Each school is represented by its own quality index \( Q \), and the quality index is a function of observable school characteristics such as school size, per-pupil teachers’ salary and per-pupil school development funds, etc. We can then write \( Q_j = Q(X_j) \), where \( Q_j \) is the quality index of school \( j \) and \( X_j \) is a vector of exogenous determinants of the quality index. The rational pupil chooses the school with the highest quality index subject to his own constraint functions. This leads us to write \( Q_i = D(Z_i) \), where \( Q_i \) is the quality index of the school chosen by individual \( i \) and \( Z_i \) is a vector of exogenous determinants of his demand. In turn, we can write that \( Q_{ij} = D_{ij}(Z_i) \) is the quality index of school \( j \) chosen by individual \( i \). In equilibrium, \( D_{ij}(Z_i) = Q(X_j) \).

Next, we use Hotelling’s canonical correlation (1935 & 1936) to estimate the coefficients of the quality index \( Q(X_j) \) and the demand for school \( j \) by individual \( i \) \( D_{ij}(Z_i) \). [See also Vinod (1968, 1969 & 1976) for a regression-like interpretation for canonical weights].

### 7.2 Model

#### 7.2.1 Demand for Schools

In this chapter we assume that the children between 5 and 20 years of age (or their guardians) first decide whether or not to go to school and then choose a school. The first decision will not be discussed in this chapter. This chapter begins by assuming that children in the relevant age group have decided to attend school. The problem they now have is to choose a school. In the remaining sections of this chapter we develop a theoretical framework explaining the choice of school and subsequently we estimate the implied demand for school function.

The maintained hypothesis of this chapter is that schools are different from each other on the quality of their service, such that some schools produce more human capital in a given
time period (for example, a year) than other schools do, and the quality measure is continuous. We further assume that school quality has many dimensions, of which in practice we observe only a few. However, school quality is a function of school inputs, which are both measurable and observable. A child chooses the school quality to maximise his utility function. The utility function is assumed to be in the following form:

$$U_i = U(Q_{ij}, \lambda_i, G_i)$$  \hspace{1cm} (7.1)

where $Q_{ij}$ is the quality of education received by individual $i$ who chooses school $j$. One of the phenomenal features of individual behaviour concerning education in Sri Lanka is that many pupils receive extra lessons from private tutors after school hours, on week-ends and during school vacations. Our survey shows that more than 30 percent of respondents receive private tuition education and on average, over one percent of the family’s per-month income is spent per-child on this purpose [see Table 5.10]. The heavily competitive education system and more importantly, the general dissatisfaction of pupils (their guardians) regarding the school education are considered by many as the primary cause for this behaviour [see Economic Review, May/June (1994)]. Therefore, $Q$ should be defined to include both the quality of the public school where the respondent attends and the quality of private tuition education.

For the ease of simplicity, one may assume that the quality of education, $Q$ is the sum of the quality of public school, $Q_s$ and the quality of private tuition education $Q_p$. $G$, is the all-other-goods, and $\lambda$ is a shift-parameter representing the preference for quality. We further assume that the shift parameter, $\lambda$ is a function of the academic environment of the family and that of the community where the respondent lives. More precisely, we assume that more educated parents and more educated communities assign a higher weight to the quality of school. In other words, they have higher $\lambda$ as compared to the $\lambda$ of the average respondent. The academic environment of the family is assumed to be represented by mother’s education (M)\(^1\) and the academic background of the community is assumed to be represented by the standardised-score (z-score) of the length of schooling of the adults (over 30 years old) living in the community. $(Z_A)\(^2\). This leads us to write $\lambda=\lambda(M,Z_A)$.

The utility function is maximised subject to the following budget constraint: $Y = G + E(Q)$, where the price of $G$ is normalised to one, $E(Q)$ is the total private expenditure on the quality of education, and $Y$ is the family income. Total private expenditure on quality of education, $E(Q)$ is the sum of $E_s(Q_s)$, $E_p(Q_p)$ and $E_d(D(Q_s))$. Where $E_s(Q_s)$ is the private expenditure on the quality of public education, $E_p(Q_p)$ is the private expenditure on the quality of private tuition education, and $E_d(D(Q_s))$ is the private expenditure on travelling to public schools. This expenditure relation assumes that travel expenditure is a positive

---

1 For consistency we choose mother’s education instead of father’s or the average education of parents. Throughout this book we use mother’s education to represent the family education background. In fact, we did estimate the model by replacing mother’s education with father’s education, but observed that the results were not sensitive to this change.

2 The average length of schooling of the adults in our sample is 8.08 years with the standard deviation of 1.15.
function of distance travelled for education. Maximising the utility function subject to the budget constraint yields the following demand functions for \( Q_s \) and \( Q_p \):

\[
Q_s = Q_p \left[ Y , \frac{\partial E_s}{\partial Q_s} , \frac{\partial E_d}{\partial D} , \frac{\partial D}{\partial Q_s} , \frac{\partial E_p}{\partial Q_p} , M, Z \right] \quad (7.2.1)
\]

\[
Q_p = Q_p \left[ Y , \frac{\partial E_s}{\partial Q_s} , \frac{\partial E_d}{\partial D} , \frac{\partial D}{\partial Q_s} , \frac{\partial E_p}{\partial Q_p} , M, Z \right] \quad (7.2.2)
\]

and therefore,

\[
Q = Q \left[ Y , \frac{\partial E_s}{\partial Q_s} , \frac{\partial E_d}{\partial D} , \frac{\partial D}{\partial Q_s} , \frac{\partial E_p}{\partial Q_p} , M, Z \right] \quad (7.2.3)
\]

Note that under standard simplifying assumptions, \( \frac{\partial E_s}{\partial Q_s} = p_s \), \( \frac{\partial E_p}{\partial Q_p} = p_p \), \( \frac{\partial E_d}{\partial D} = p_d \) and \( \frac{\partial D}{\partial Q_s} = d_Q \), where \( p_s \) is the unit price of the quality of education at public schools, \( p_p \) is the unit price of the quality of education at private tuition schools, \( p_d \) is the price per-one kilometre distance travelled for schooling, and \( d_Q \) is the distance needed per-unit increase of education quality.

In this study we assume that \( p_s \) and \( p_p \) are constant for all individuals for a given time and \( p_d \) and \( d_Q \) are positive functions of the distance travelled, \( D \).

We can write the following cross-sectional demand for quality of education function as

\[
Q = Q \left[ Y , M, Z, D \right] \quad (7.3)
\]

where all the variables are already defined\(^3\). Assuming a semi-log functional form, and allowing for non-linearity of the effect of \( Y \) and \( D \) and interactions between \( M, Y \) and \( D \), we can write the following demand for school quality function as

\[
Q_i = \exp \left( \beta_0 + \beta_1 Y_i + \beta_2 Y_i^2 + \beta_3 D_i + \beta_4 D_i^2 + \beta_5 M_i + \beta_6 (MY) + \beta_7 (MD) + \varepsilon_i \right) \quad (7.3.1)
\]

### 7.2.2 Determinants of School Quality Index

By following Chapter 4 (equation 4.6), we write the school quality index in the following form:

\[
Q = \gamma \left( \frac{G}{T} \right)^{\gamma_1} N^{-\gamma_2} \left[ D K^{-\gamma_p} + (1-\delta) L^{-\gamma_p} \right]^{\gamma_3} e^{\varepsilon_2} \quad (7.4)
\]

This is the model we estimated in Chapter 4. In this chapter we have an additional problem, i.e. inclusion of \( \ln(G/T) \) into the production function specification was not possible due to high co-linearity. Therefore, \( \ln(G/T) \) was also excluded from the specification at the estimation stage.

---

\(^3\) Equation (7.3) defines \( Q \) and the sum of \( Q_s \) and \( Q_p \). However, due to data limitations, we are restricted only to \( Q_s \) in this chapter.
We can now substitute the right-hand side of equation (7.4) into the left-hand side of equation (7.3.1) and re-write

\[
\gamma_o \left( \frac{G}{T} \right) \gamma_1 N^{-\eta} \left[ \delta K^{-\nu} + (1-\delta) L^{-\nu} \right]^{-\nu} e^{\nu - 2} = 
\exp(\beta_o + \beta_1 Y_1 + \beta_2 Y_1^2 + \beta_3 D_1 + \beta_4 D_1^2 + \beta_5 M_1 + \beta_6 Z_1 + \beta_7 (MY)_1 + \beta_8 (MD)_1 + \epsilon_1)
\]

(7.5)

where all the variables and coefficients are already defined.

7.3 Estimation Procedure
7.3.1 An Overview of the Estimation Procedure

It is clear that equation (7.5) cannot be estimated by the OLS. One method available in the literature is canonical regression. This idea is an extension of Hotelling’s Canonical Correlation. [See Hotelling (1935) and (1936) for the original work.] Canonical Correlation is the simple Pearson’s Correlation between two linear combinations of variables. In other words, it is the simple correlation between two unobserved composite indices.

The canonical regression technique is applicable only if the relations on both sides of the equation are intrinsically linear. (They are linear or can be converted into linear equations.) It is apparent that by taking natural logarithms we can convert the demand function into a linear relationship. However, the production function cannot be converted. Therefore, we use the second-degree Taylor’s expansion of the production function (See Appendix 4A). Linear versions of both equation (7.3.1) and (7.4.1) are given below:

\[
\ln Q_i = \beta_o + \beta_1 Y_1 + \beta_2 Y_1^2 + \beta_3 D_1 + \beta_4 D_1^2 + \beta_5 M_1 + \beta_6 Z_1 + \beta_7 (MY)_1 + \beta_8 (MD)_1 + \epsilon_1 
\]

(7.6)

\[
\ln Q = \ln \gamma_o + \gamma_1 \ln (G/T) - \eta \nu \ln N + \nu \delta \ln K + \nu (1-\delta) \ln L - \frac{1}{2} \rho \nu \delta (1-\delta) [\ln K - \ln L]^2 + \epsilon_2 
\]

(7.7)

7.3.2 Canonical Correlation
7.3.2.1 The Technique

Suppose we have two matrices Y and X: Y with k variables and X with g variables arranged by rows, respectively^4. Each contains N observations. Further assume that all variables are standardised. We define two linear composites U and V: U = \alpha Y and V = \beta X. The canonical correlation is then defined as the simple Pearson’s correlation between U and V. Here, the methodology is to find \alpha and \beta such that the correlation between U and V is maximised. Since all the variables in Y and X matrices are standardised, U and V are also standardised and therefore the correlation between U and V is the covariance between the two. We can define the covariance between U and V, Cov(U,V) = E(UV) = E(\alpha Y \beta) = \alpha E(YX) \beta = \alpha R_{XY} \beta. R is the correlation matrix between Y and X variables (inter-set-

^4 Throughout we assume that g is greater than k.
inter-correlation). The maximisation is done subject to two constraints, namely both U and V have unit variance. This yields the following maximisation problem:

\[ \varphi = \alpha^\prime R_{YX} \beta - \lambda (\alpha^\prime R_{YY} \alpha - I) - \mu (\beta^\prime R_{XX} \beta - I) \]  

(7.8)

where \( \lambda \) and \( \mu \) are Lagrange multipliers and the expressions in the two constraints are the variances of U and V respectively. In order to have a non-trivial solution for this problem, the first-order condition requires \[ |R_{YX}^{-1} R_{YY} R_{XX}^{-1} R_{XY} - \lambda I| = 0 \]

where all Rs are matrices of relevant correlation coefficients as indicated by subscripts. The operator \(|...|\) indicates the determinant value. Since this equation will have more than one solution for \( \lambda \), we have the subscript \( j \) with \( \lambda \). It is very easy to see that these solutions to \( \lambda \) are the characteristic roots of the determinant \[ |R_{YX}^{-1} R_{YY} R_{XX}^{-1} R_{XY} - \lambda_j I| \]

and therefore of equation (7.8). The number of distinct characteristic roots is equal to the number of variables in the smaller matrix, in this case \( k \). We can also show that the square root of a characteristic root gives a canonical correlation [for details, see Anderson (1984) and Vinod (1968)].

Canonical correlation is mostly used in psychology where the researchers’ interest is to measure the association between pairs of complex phenomena such as mental and physical fitness, ability and the various determinants of it. [See inter-alia Morrison (1967) and Thompson (1984) for more illustrative examples]. Economists’ interest in this method is rather recent. [See Vinod (1968), (1976), Kuylen and Verhallen (1981) and Rijken Van Olst (1981) for early economic applications and Chizmar and Zak (1983), (1984) and Chizmar and McCarney (1984), Cohn et al. (1989) and Gyimah and Gyapong (1991) for recent applications].

Vinod (1968) further develops this idea and gives a regression-like interpretation to canonical weights (vectors \( \alpha \) and \( \beta \) above). A comprehensive discussion of the econometric applications of canonical regression is also available in Rijken Van Olst (1981). [See Chetty (1969), Dhrymes et al. (1969) and Rao (1969) for criticism of Vinod (1968)]. The following paragraphs briefly discuss Vinod’s canonical regression.

According to the above description of canonical correlation, we can write U= AY, V= BX and because U and V are standardised variables (by assumption), we can write U = RV or AY = RBX, where A and B re-arrange vectors \( \alpha \) and \( \beta \) above into diagonal matrices. R is a diagonal matrix of canonical correlation coefficients. This provides the following estimates of canonical regression coefficients for each canonical correlation coefficient.

\[ \hat{\alpha}_m = \hat{\alpha}_m \]  

(7.9.1)

\[ \hat{\beta}_h = \hat{r}_j \hat{\beta}_h \]  

(7.9.2)

where \( \hat{\alpha} \) and \( \hat{\beta} \) are the estimates of canonical regression coefficients, \( \hat{\alpha} \) and \( \hat{\beta} \) are the estimates of canonical weights for Y and X matrices, respectively. Subscripts denote variable names, \( \hat{r} \) is the canonical correlation and its subscript denotes the order of the canonical correlation coefficient. For example, \( r_j \) is the \( j^{th} \) canonical correlation coefficient.
Subscripts \( m \) and \( h \) denote the canonical weights attached with the \( m \)th variable in matrix \( Y \) and \( h \)th variable in matrix \( X \), respectively.

### 7.3.2.2 Tests for Statistical Validity

There are several tests for the statistical validity of a canonical regression. Three such tests are described below: \( R^2 \): proportion of explained variance of \( Y \)(and \( X \)) variables; \( R^2 (Y \mid X) \) and \( R^2 (X \mid Y) \): redundancy coefficient of \( Y \) given \( X \) and the redundancy coefficient of \( X \) given \( Y \), respectively; \( V(Y \mid X) \) or \( V(X \mid Y) \): proportion of redundancy of \( Y \) given \( X \) or \( X \) given \( Y \) for a given canonical correlation coefficient.

The proportion of explained variance is conceptually similar to the \( R^2 \) in the regression context. We can estimate this measure for both \( X \) and \( Y \) variables separately. The formula is 

\[
R^2 = \frac{\sum_{j=1}^{n} (CL_j)^2}{n},
\]

where \( CL \) is the canonical loading of a given matrix. Canonical loadings are the simple partial correlation coefficients of canonical variates (\( U \) or \( V \)) with their respective matrices (\( Y \) or \( X \), respectively). In the above formula \( n \) refers to number of variables in the matrix (\( g \) and \( k \), respectively).

But this measure takes only one set of variables into consideration. Therefore, the redundancy coefficient is defined by multiplying the above coefficient by the square of the canonical correlation coefficient. This measures the explanatory power of one variable set given the explanatory power of the other set.

The proportion of redundancy is another measure. This can be estimated by dividing the redundancy coefficient for a given canonical correlation by the sum of all redundancy coefficients for the given matrix.

There are also some measures for the statistical significance of the canonical regression. For example, C.R. Rao’s F- statistic and Bartlett’s Chi-square\(^5\) are given below.

\[
\text{Rao’s F} = \frac{(1 - \Lambda_1^{1/s}) - ms + 2v}{\Lambda_1^{1/s} \text{kg}} \quad (7.10.1)
\]

\[
\text{Bartlett’s Chi-square } \chi^2 = -[(N-1) - \frac{1}{2}(k+g+1)] \ln \Lambda \quad (7.10.2)
\]

where \( N \) is number of observations, \( k \) and \( g \) are number of variables in \( Y \) and \( X \) matrices, respectively. \( \Lambda = \prod_{i=1}^{\min(k,g)} (1 - \lambda_i) \), \( m = [N - 1 - \frac{1}{2}(k + g + 1)] \), \( s = \left\lfloor \frac{k^2 g^2 - 4}{k^2 + g^2 - 5} \right\rfloor ^{1/2} \) and,

\[
v = \left\lfloor -1 \left( \frac{kg - 2}{4} \right) \right\rfloor .
\]

\(^5\) This is as reported in Timm and Carlson (1976).
Degrees of freedom for $F$ distribution are $(kg)$ for the numerator and $(ms+2v)$ for the denominator. Bartlett's Chi-square is with $kg$ degrees of freedom.

7.4 Estimation and the Results

7.4.1 Data

Data for the empirical estimation of model (7.5) has been gathered from a sample survey and the School Census of 1992. The sample survey was conducted in the fourth quarter of 1996 in Sri Lanka [see Chapter 5]. This survey provides all required variables except school characteristics. Due to the practical difficulties of conducting a countrywide survey, our sample has been selected from one particular district: the Kandy District. The school-input measures were then collected from the School Census of 1992. These data have been described in Chapter 5.

7.4.1.1 Descriptive Statistics

Descriptive statistics of the data used in the analysis are reported in Table 7.1 below. Since the analysis in this chapter is based only on the current enrollees, figures in Table 7.1 were calculated by using the information only from current enrollees. There are 1,014 full-time students in our sample. Therefore, the figures reported in this table are different from the descriptive statistics reported in Chapter 6 (Table 6.1). In addition, the definition of some variables is also changed. In this chapter the economic background of the nuclear family is represented by the family income per-month, whereas parents' income is used in Chapters 3 and 6. In Chapter 3 we argued that the current family income does not properly represent the economic status of school leavers at the time of leaving school. Under those circumstances, the current income of parents may be more representative of the economic status of nuclear families of school leavers, if we can assume that parents' economic conditions are more stable over time. We do not have that problem in this chapter because all the respondents here are current enrollees. And finally, all the school characteristics are in natural log form.
Table 7.1
Descriptive Statistics of Variables Used in Canonical Regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables in Production Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Number of Pupils: InN</td>
<td>6.74</td>
<td>0.90</td>
<td>3.00</td>
<td>8.08</td>
</tr>
<tr>
<td>Log Extra School Income: lnK</td>
<td>10.04</td>
<td>3.82</td>
<td>-6.91</td>
<td>15.20</td>
</tr>
<tr>
<td>Log Teacher Salary: lnL</td>
<td>11.41</td>
<td>0.90</td>
<td>8.57</td>
<td>12.82</td>
</tr>
<tr>
<td><strong>Variables in Demand Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per-member Family Income (Rs.000) ; Y</td>
<td>0.83</td>
<td>0.81</td>
<td>0.06</td>
<td>10.25</td>
</tr>
<tr>
<td>Mother’s Education (years): M</td>
<td>7.98</td>
<td>3.82</td>
<td>0.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Length of Schooling of Adults (years)</td>
<td>8.08</td>
<td>1.15</td>
<td>0.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Standardised-score of Adults Schooling</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.84</td>
<td>1.22</td>
</tr>
<tr>
<td>Distance Travelled to the School (k.m.): D</td>
<td>2.65</td>
<td>4.25</td>
<td>0.00</td>
<td>36.70</td>
</tr>
<tr>
<td><strong>Variables in Cost Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private Expenditure for Education</td>
<td>206.31</td>
<td>207.67</td>
<td>19.68</td>
<td>1730</td>
</tr>
<tr>
<td>Expenditure for Transport &amp; Boarding</td>
<td>39.31</td>
<td>97.05</td>
<td>0.00</td>
<td>1000</td>
</tr>
<tr>
<td>Expenditure for all other Educational Services</td>
<td>167.31</td>
<td>155.76</td>
<td>19.47</td>
<td>1480</td>
</tr>
</tbody>
</table>

Note: Zeros in K were replaced with 0.001 (an arbitrary-small number) in order to take natural logarithms.

Distance travelled is the distance from the usual residence to their school.

Most of the variables in Table 7.1 have already been described in Chapter 5 and will therefore not be repeated here. Per-member family income is on a monthly basis. This is the per-member income in thousand rupees received by all family members during the reference period. The reference period again varies for different income items. For example, income received from agricultural activities refers to the last season, whereas for all other income sources the reference period is the previous month.

Table 7.2
Correlation Matrix

<table>
<thead>
<tr>
<th></th>
<th>InN</th>
<th>lnK</th>
<th>lnL</th>
<th>lnK-lnL</th>
<th>TC</th>
<th>Y</th>
<th>Y²</th>
<th>M</th>
<th>MY</th>
<th>MD</th>
<th>Z</th>
<th>D</th>
<th>D²</th>
</tr>
</thead>
<tbody>
<tr>
<td>InN</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lnK</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td>-0.85</td>
<td>0.44</td>
<td>0.56</td>
<td>0.37</td>
<td>0.39</td>
<td>0.39</td>
<td>0.36</td>
<td>0.16</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>lnL</td>
<td>0.93</td>
<td>0.80</td>
<td>1.00</td>
<td>-0.46</td>
<td>1.00</td>
<td>0.56</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.23</td>
<td>0.26</td>
<td>1.00</td>
</tr>
<tr>
<td>lnK-lnL</td>
<td>-0.30</td>
<td>-0.50</td>
<td>-0.46</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC</td>
<td>0.41</td>
<td>0.34</td>
<td>0.37</td>
<td>-0.10</td>
<td>1.00</td>
<td>0.56</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Y</td>
<td>0.34</td>
<td>0.29</td>
<td>0.35</td>
<td>-0.09</td>
<td>0.56</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Y²</td>
<td>0.19</td>
<td>0.17</td>
<td>0.20</td>
<td>-0.05</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>M</td>
<td>0.36</td>
<td>0.39</td>
<td>0.42</td>
<td>-0.23</td>
<td>0.39</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>MY</td>
<td>0.38</td>
<td>0.34</td>
<td>0.40</td>
<td>-0.12</td>
<td>0.39</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>MD</td>
<td>0.15</td>
<td>0.17</td>
<td>0.17</td>
<td>-0.07</td>
<td>0.36</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Z</td>
<td>0.30</td>
<td>0.36</td>
<td>0.41</td>
<td>-0.21</td>
<td>0.36</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.34</td>
<td>0.16</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>D</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.23</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>D²</td>
<td>-0.08</td>
<td>-0.04</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.16</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Note: TC= Total Private Cost of Education. Definition is given in Chapter 5, section 5.5.3.

Table 7.2 reports the correlation matrix of the variables in the model. The correlation matrix is divided into four quadrants. The top left quadrant reports the partial correlation

---

6 An alternative might be to apply family size equivalence scales. Without robust equivalence scales estimated for Sri Lanka however, this would only add an element of arbitrariness to the model. [For details on equivalence scales, see Folkersima (1996) inter-alia].
coefficients between the variables in the production function and the bottom left quadrant shows the correlation coefficients between the variables in the production function and those in the demand function. In the literature this is called the intersect-inter-correlation matrix. The bottom right quadrant reports the correlation coefficients between the variables in the demand function.

The correlation matrix shows that the variables in the production function are highly correlated. The presence of high multicollinearity has to be resolved before estimating the model. Vinod (1976) suggested applying the Canonical Ridge estimator to overcome the problem. However, as is well known, ridge regression results are not consistent. The only advantage of ridge regression is that it is efficient. Furthermore, the application of the Canonical Ridge Regression is rather cumbersome. However, with the OLS or Canonical regressions we can still have unbiased regression estimates. Therefore, in this chapter we prefer to obtain unbiased estimates of the regression parameters although they are not efficient.

7.4.2 Regression Results

Table 7.3 below presents the estimates of the Canonical Regression of equations (7.6) and (7.7).

The R² is relatively high for the production function, thus indicating that a fair proportion of the variation of the production function is captured by our specification. Nearly 12 percent of the total variation of the variables in the demand side is captured by our specification. The proportional redundancy rate shows that over 90 percent of the total variation on each side of our specification is captured by the first canonical correlation. This justifies our decision to leave higher order canonical correlation coefficients unexplained. Chi-squared and Rao’s F statistics show that the specification is statistically significant at any reasonable level of significance.

We can use the regression results reported in Table 7.3 to calculate the parameters of the CES production function and various elasticity coefficients of demand for quality. According to equation (7.5), the top part of Table 7.3 provides estimates of -ηv, vδ, v(1-δ) and $-\frac{1}{2}v\delta(1-\delta)$, respectively. We can solve the reported coefficients for each of these coefficients. Table 7.4 reports the calculations.

---

7 Note that intercepts are not calculated. Estimation was done with mean deviations of all variables and therefore intercepts are zero.
Table 7.3
Canonical Regression Estimates: First Canonical Correlation

<table>
<thead>
<tr>
<th></th>
<th>Production</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnN</td>
<td>-1.011</td>
<td>0.154</td>
</tr>
<tr>
<td>lnK</td>
<td>0.431</td>
<td>-0.049</td>
</tr>
<tr>
<td>lnL</td>
<td>1.090</td>
<td>0.053</td>
</tr>
<tr>
<td>([\lnK - \lnL]^2)</td>
<td>0.027</td>
<td>0.003</td>
</tr>
<tr>
<td>(R_{2}^2)</td>
<td>0.506</td>
<td>0.309</td>
</tr>
<tr>
<td>Redundancy</td>
<td>0.199</td>
<td>0.122</td>
</tr>
<tr>
<td>Prop. Redundancy</td>
<td>0.931</td>
<td>0.926</td>
</tr>
</tbody>
</table>

Note: Coefficients reported in the lower part of the table are the products of canonical weights and the first canonical correlation.

t: t statistics for zero null hypothesis.
Loadings: Partial correlation coefficients between the predicted composite indices and their respective variables.

Table 7.4
Estimates of the Coefficients of CES Production Function

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\nu)</td>
<td>1.52</td>
</tr>
<tr>
<td>Distribution Parameter: (\delta)</td>
<td>0.28</td>
</tr>
<tr>
<td>(\eta)</td>
<td>0.67</td>
</tr>
<tr>
<td>(\rho)</td>
<td>-0.17</td>
</tr>
<tr>
<td>Size-elasticity: (-\eta\nu)</td>
<td>-1.01</td>
</tr>
<tr>
<td>Return to Scale: (\nu(1-\eta))</td>
<td>0.51</td>
</tr>
<tr>
<td>Elasticity of Substitution: (\sigma)</td>
<td>1.21</td>
</tr>
</tbody>
</table>

For further information on the definitions of these coefficients, see section 4.4.2 of Chapter 4.

A comparison of the results in this chapter with the results reported in Chapter 4 (Table 4.6) shows some dissimilarity. The school output measure in Chapter 4 was the odds-ratio of passing the OL, whereas in this chapter the output is more general and unobserved. The samples used in the two analyses and the statistical techniques are also different. Some determinants of the school production function in Chapter 4 (G/T) are removed from the
analysis in this chapter. The observed dissimilarities of the parameters of the production function may be explained in terms of these differences. Above all, what is most compelling is that both findings lead us to the same policy recommendation.

The **size-elasticity** is negative and substantial, thus indicating that schools are overcrowded. This was negative but rather inelastic in Chapter 4. Return to scale for per-pupil inputs (v) is greater than one. Schools are experiencing an *increasing return to scale*. A 1 percent increase of all the inputs per-pupil associates with a more than 1 percent improvement of the quality of schools. In other words, to expand schools by adding more teachers and other facilities, while keeping number of pupils constant, may increase the quality of education. This is not consistent with the findings reported in Chapter 4, where we found decreasing returns to per-pupil inputs as well. Estimates of elasticity of substitution (σ) and distribution parameter (δ) are quite similar to their estimates in Chapter 4. These coefficients are interpreted in Chapter 4.

Table 7.5
Demand Elasticity of School Quality

<table>
<thead>
<tr>
<th>Variable</th>
<th>Elasticity Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y : Family Per-member Income (in Rs.)</td>
<td>0.27</td>
</tr>
<tr>
<td>M : Mother's Education</td>
<td>0.47</td>
</tr>
<tr>
<td>D : Distance Travelled</td>
<td>0.17</td>
</tr>
<tr>
<td>Z: Community Education a</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Note: *The average Z is zero by definition. Therefore, in this cell we report the regression coefficient itself. This measures the effect of increasing the community education by one standard deviation unit.*

The demand function specified by equation (7.3.1) suggests non-constant elasticity measures. Our estimates show that the demand for quality of schooling is inelastic to all the variables in the demand function. Community and family education background seem to be the most important determinants of the demand for quality of schooling.

The income elasticity first increases with income and then declines. A positive interaction between income and mother’s education shows that the income elasticity is higher for children with a higher academic environment at home. Interaction between income and community education background was also tested and it turned out to be statistically insignificant. Variation of the income elasticity with income and mother’s level of education is depicted by Figure 7.1 below. The figure has four projections. Each represents the income-elasticity for different levels of mother’s education. The lowest line represents the income elasticity when the mother’s education is zero. The next line represents income elasticity when the mother’s education is at average. The third line is the income elasticity for mode of the mother’s education, and the top line is associated with the situation where mother has a degree or above. The distance elasticity has the same pattern.
Fig. 7.1: Simulation of Income Elasticity of Quality of Education: by Mother's Education

Note: $E_1$ is the path of Income Elasticity when Mother’s Education is at Maximum
$E_2$ is the path of Income Elasticity when Mother’s Education is at Mode
$E_3$ is the path of Income Elasticity when Mother’s Education is at Mean
$E_4$ is the path of Income Elasticity when Mother’s Education is at Minimum

The hump-shaped income elasticities may be explained in terms of Becker and Tomes (1986). Becker and Tomes (1986) argued that altruistic parents invest in their children in two different forms: in human capital and in non-human capital. For example, they invest in the education of their children and they also bequest their physical inheritance to the children. The marginal rate of return to the physical capital is constant. It is a given rate. However, the marginal rate of return to human capital is diminishing with the amount of investment. This is because there is a maximum that parents’ investment can do in the human capital production function. Once the maximum is reached, further investment adds nothing to the human capital stock of children. As long as the marginal return to investment in human capital is greater than that of non-human capital, a 1 percent increase in the family income will associate with an increase of the investment in human capital. However, there is a critical value of investment in human capital, where the marginal return to human capital equalises the marginal return to non-human capital. Once the critical value is reached parents will no longer increase the investment in human capital. In general, it is more likely that rich parents are closer to the critical value of human capital investment than poor parents. Assuming that the choice of school quality is a form of parental investment in the human capital formation of their children, we can apply the above theory to explain the variation of income elasticity depicted in Figure 7.1 above. The positive effect of mother’s education on the income elasticity may be explained either in terms of an ability effect or in terms of non-monetary values associated with education. The positive effect of mother’s education on the ability of offspring will predict that the offspring of educated mothers have a higher marginal rate of return to the quality of education, and they need a large amount of investment to reach the critical point.
7.5 Cost of Education Quality

We have gathered information on various types of cost items in the sample survey. Definitions and some descriptive statistics of these cost items are presented in Chapter 5. As we described in Chapter 5, these cost items represent different types of costs related with education. Remember that our sample covers only the public schools and therefore what students pay to their schools is not a substantial contribution. Our estimates show that direct payments to schools in terms of facility fees and donations account for only 4 percent of the total private expenses of education.

Presumably, the private cost of education is positively correlated with the quality of education. In this section we regress the private cost of education on the quality of schools estimated by using the canonical regression. Table 7.6 reports two different estimates of the cost function: one with the quality predicted from the demand function, and the other with the quality predicted from the production function reported in Table 7.3.

Table 7.6

<table>
<thead>
<tr>
<th></th>
<th>Q: Demand</th>
<th>Q: Production</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>t-stat</td>
</tr>
<tr>
<td>Predicted Quality (power 3)</td>
<td>1.78</td>
<td>4.31</td>
</tr>
<tr>
<td>Male Respondents</td>
<td>-3.36</td>
<td>-0.24</td>
</tr>
<tr>
<td>Urban Residence</td>
<td>-13.38</td>
<td>-0.76</td>
</tr>
<tr>
<td>Ethnic Sinhalees</td>
<td>-31.23</td>
<td>-1.93</td>
</tr>
<tr>
<td>Quality (power 3) &amp; Male</td>
<td>0.23</td>
<td>0.87</td>
</tr>
<tr>
<td>Quality (power 3) &amp; Urban</td>
<td>0.73</td>
<td>2.42</td>
</tr>
<tr>
<td>Quality (power 3) &amp; Ethnic Sinhala</td>
<td>1.37</td>
<td>3.74</td>
</tr>
<tr>
<td>Constant</td>
<td>112</td>
<td>6.91</td>
</tr>
<tr>
<td>Adjusted R squared</td>
<td>0.42</td>
<td>0.33</td>
</tr>
<tr>
<td>F (7, 1006)</td>
<td>107</td>
<td>73</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>1014</td>
<td>1014</td>
</tr>
</tbody>
</table>

Note: School Quality is predicted using the demand and production functions reported in Table 7.3.
Q: Demand: Quality is predicted using Demand function.
Q: Production: Quality is predicted using Production function.

The reported cost functions in Table 7.6 are cubic equations. We estimated linear, second and third degree polynomial specifications of the cost function. When the third degree term is included, school quality and its square turned out to be insignificant. Adding first and second-degree terms to the cubic equation does not increase its explanatory power. For example, the adjusted R-square of the cubic cost function (without dummies) is 0.41 and it increases only to 0.42 when the first and second-degree terms are included. Therefore, the final version of the cost function is cubic. We further add three dummies and their interactions with quality to allow total, average and marginal cost to vary over gender, place of residence and ethnicity.

Regression coefficients are highly sensitive to the choice of the quality-measure. In general, gender does not have any effect on the average or the marginal cost of quality of
schooling. Urban residents and ethnic Sinhalees experience lower average costs and higher marginal cost schedules.

In fact, the analysis reported in this section is preliminary and requires improvement, both in theoretical and in technical terms. We will not pursue this any further.

7.6 Summary, Conclusions and Recommendations

7.6.1 Summary & Conclusions

In this chapter we used the canonical regression to estimate the equilibrium condition of the market for school quality. One of the advantages of this method is that it estimates both the production and the demand for school quality simultaneously. Despite the inherent limitations of the data set, we observe that our method draws conclusions similar to previous studies using different statistical techniques.

Estimates of the parameters of the production function show us that the entire production process is dominated by a teacher input intensive production technology. This is what we already found in Chapter 4.

Schools in our sample are experiencing an increasing return to scale for per-pupil inputs. The size-elasticity is minus one, indicating that the average public school in Sri Lanka is overcrowded. A 1 percent increase of number of pupils is associated with a 1 percent decrease of the quality of education. Therefore, the expansion of schools (increase both pupils and inputs) reduces the quality of service. Most of the findings in this chapter with regard to the school production function are similar to the findings in Chapter 4.

The demand function to which we have paid more attention in this chapter reveals an interesting story, which is also very useful for policy making purposes. Demand for school is more sensitive to family and community education than to family income and distance. Estimates of the demand elasticities show that the academic environment of the family (M) and that of the community where the respondent lives (Z) have the highest impact on the demand for quality of schooling. Figure 7.1 further shows that income elasticity is also positively correlated with the academic environment of the family. The same is true for the distance elasticity. In fact, this has a very strong policy implication. According to our model, M and Z represent the given factors for policy makers: family and social background of the pupils. Policy makers can intervene only through Y: financial compensations. However, our results show that the effects of financial compensation are positively correlated with family and social background of pupils. This limits the policy makers’ capacity to resolve the problem by providing financial aid, because for whichever the reason, pupils from low education backgrounds respond slowly.
Chapter 7: Production, Demand & Private Costs

7.6.2 Recommendations

The conclusions of this chapter are limited by several drawbacks of the methodology and data availability. A sampling bias, particularly with respect to the estimation of the production function, is to be expected. Our argument is the following. We have chosen a sample on a household basis. The sample is stratified by census blocks [see Chapter 5]. Therefore, our sampling method is more likely to choose a greater number of individuals from large schools located in the vicinity of the residence in, for example, 20 percent of respondents from 10 schools (see Chapter 5, section 5.5.5). This essentially suggests that the results of the school production function have to be verified by an alternative estimate of the same model with a more representative sample. More careful sample design, which picks up both households as well as schools on a random basis into the sampling frame, will also be a solution to this problem. This study has been based entirely on a sample selected from one particular district, which suggests an expansion of the sample to represent the entire country.

The analysis in this chapter was restricted due to data and technical problems. For example, the quality of school administration, an important determinant of the quality of school, was not available to us. We also had to drop the percentage of graduate teachers from the model because of a high co-linearity problem. The percentage of graduate teachers represents the education composition of teachers, which presumably increases the efficiency of the school production process. A detailed discussion of this issue was available in Chapter 4.