Barro's fertility equations: the robustness of the role of female education and income

PAULA K. LORGELLY, STEPHEN KNOWLES and P. DORIAN OWEN*

Department of Economics, University of Otago, PO Box 56, Dunedin, New Zealand

Barro and Lee (1994) and Barro and Sala-i-Martin (1995) find that real per-capita GDP and both male and female education have important effects on fertility in their cross-country empirical studies. In order to assess the robustness of their results, their estimated models are subjected to specification and diagnostic testing, the effects on the model of using the improved Barro and Lee (1996) cross-country data on educational attainment of the population aged 15 and over are examined, and the different specifications used by Barro and Lee and by Barro and Sala-i-Martin compared. The results obtained suggest that their fertility equations do not perform well in terms of diagnostic testing, and are very sensitive to the use of different vintages of the educational attainment proxies and of the Summers-Heston cross-country income data. A robust explanation of fertility, to link with empirical growth equations, has, therefore, not yet been found; further work is required in this area.

I. INTRODUCTION

An interesting aspect of the effect of human capital accumulation on economic growth is the potentially different roles of male and female education. Female participation rates in the paid workforce are generally lower than for males, but female education is widely believed to produce important social gains and to have indirect effects on measured productivity growth (e.g. Subbarao and Raney, 1995). Robert Barro and his colleagues have produced a series of cross-country empirical growth studies, including Barro and Lee (1994) (hereafter B–L) and Barro and Sala-i-Martin (1995) (hereafter B–SM), that incorporate gender-specific education variables in their estimated models.[1] B–L and B–SM also estimate regression equations explaining fertility, infant mortality, life expectancy and school enrolments, with female education among the explanatory variables. Part of their motivation appears to be a desire to examine the indirect effect of female education on growth via its effects on fertility, health status and school enrolments. The results from their estimated growth equations 'are somewhat disappointing in terms of demonstrating an important role for educational attainment in the growth process' (B–L, 1994: 32). By contrast, female education appears to have a more important influence on fertility, the health indicators and schooling.[2] In particular, fertility is inversely related to female educational attainment.

The objective of this study is to examine the robustness of the B–L and B–SM results on the determinants of the fertility rate, particularly the role of female education, in the context of their chosen models. Although relatively neglected in applied work in economics, there are compelling arguments as to why replication can be very useful in improving the quality and reliability of empirical results

* Corresponding author. E-mail: dowen@commerce.otago.ac.nz

[1] B–L (1994) and B–SM (1995) find that female schooling has what they describe as a 'puzzling' (marginally significant) negative impact on economic growth.

[2] B–L (1994: 43) conclude that female schooling 'has more important [than in the growth regressions] influences . . . on choices of the quantity [measured by the fertility rate] and quality [measured by schooling, infant mortality and life expectancy at birth] of children, effects that should impact on growth in the long run'.

(e.g. Dewald *et al.*, 1986; Tomek, 1993). In terms of Tomek's (1993) terminology, both 'confirmation' (estimating the original models using the original data) and 'replication' (estimating the original models using new data) are examined.

The sensitivity of the B–L and B–SM results to the use of B–L's (1996) updated cross-country data on educational attainment are considered. The new data set is believed to be superior to the B–L (1993) data used by B–L (1994) and B–SM (1995). B–L (1996: 222) express interest in using the data to further investigate the 'impact of adult schooling on fertility, child mortality, and education of children'. However, to date, estimates of their fertility model using the new educational attainment data have not appeared in the literature.[3]

The reasons are also examined for the relatively major changes in the specification of the fertility equation in B–SM (1995) compared to the earlier B–L (1994) formulation. The most striking change is the omission of life expectancy and infant mortality, both of which B–L find to be significantly correlated with fertility.

Other studies have examined the determinants of fertility using cross-country data, e.g. Subbarao and Raney (1995) and Schultz (1994). However, B–L and B–SM's work on modelling fertility is of particular interest because of its emphasis on the linkages between the fertility and growth equations. Also, Barro and his colleagues' contributions are particularly influential and have been credited as having 'shaped the research agenda in growth empirics' (Caselli *et al.*, 1996: 364).[4]

Sections II and III re-estimate, respectively, the B–L and B–SM fertility equations, and subject them to specification and diagnostic testing. Section IV attempts to discover why the specification of the B–SM fertility equation is different from that of B–L. In Section V, both models are re-estimated using B–L's (1996) new cross-country educational data. Section VI concludes.

II. THE BARRO–LEE MODEL

B–L (1994) estimate their fertility model in a seemingly unrelated regression equations (SURE) system, utilizing data for two years (1965 and 1985).[5,6] B–L's fertility equations can be represented as:

$$\begin{aligned}
\ln(\mathrm{FERT})_{1j} = \beta_{10} &+ \beta_{11}\ln(\mathrm{GDP}_{1j}) + \beta_{12}(\ln(\mathrm{GDP}_{1j}))^2 \\
&+ \beta_{13}\mathrm{FEMALE}_{1j} + \beta_{14}(\mathrm{FEMALE}_{1j})^2 \\
&+ \beta_{15}\mathrm{MALE}_{1j} + \beta_{16}(\mathrm{MALE}_{1j})^2 \\
&+ \beta_{17}\ln(\mathrm{LE}_{1j}) + \beta_{18}(\ln(\mathrm{LE}_{1j}))^2 \\
&+ \beta_{19}\mathrm{MORT}_{1j} + \beta_{1,10}(\mathrm{MORT}_{1j})^2 + \varepsilon_{1j} \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\ln(\mathrm{FERT})_{2j} = \beta_{20} &+ \beta_{21}\ln(\mathrm{GDP}_{2j}) + \beta_{22}(\ln(\mathrm{GDP}_{2j}))^2 \\
&+ \beta_{23}\mathrm{FEMALE}_{2j} + \beta_{24}(\mathrm{FEMALE}_{2j})^2 \\
&+ \beta_{25}\mathrm{MALE}_{2j} + \beta_{26}(\mathrm{MALE}_{2j})^2 \\
&+ \beta_{27}\ln(\mathrm{LE}_{2j}) + \beta_{28}(\ln(\mathrm{LE}_{2j}))^2 \\
&+ \beta_{29}\mathrm{MORT}_{2j} + \beta_{2,10}(\mathrm{MORT}_{2j})^2 + \varepsilon_{2j} \qquad (2)
\end{aligned}$$

where FERT is the total fertility rate (children per woman).[7] GDP is real per-capita GDP. FEMALE and MALE are, respectively, average years of schooling of the female and male population aged 25 and over, LE is life expectancy at birth, and MORT is the infant mortality rate. Subscripts ij refer to country j in year i (where $i = 1$ for 1965 and $i = 2$ for 1985). The squared terms are included to allow for non-linear relationships between the explanatory variables and fertility.

The error terms are (implicitly) assumed to satisfy

$$\begin{aligned}
E(\varepsilon_{ij}) &= 0, & \forall i,j \\
E(\varepsilon_{ij}^2) &= \sigma_{ii}^2 \neq 0, & \forall i,j \\
E(\varepsilon_{1j}\varepsilon_{2j}) &= \sigma_{12} \neq 0, & \forall j \\
E(\varepsilon_{hj}\varepsilon_{ik}) &= 0, & h,i = 1,2; j \neq k
\end{aligned}$$

[3] Barro (1997) uses the new education data in estimating a growth equation and reports (p. 21, fn. 11) that '[with] the revised data on education, the estimated female coefficients are essentially zero'. This would further focus interest on the indirect effects of female education on growth, but no results are reported for the determinants of fertility, the health indicators or schooling.

[4] Fertility equations are also estimated in Barro (1991, 1992). However, in both papers, the education variable is not disaggregated by gender; also, education is found to have a positive impact on growth, so there is no puzzling negative impact of female education on economic growth to explain. In both papers, exactly the same variables appear in the fertility equations as in the respective growth equations (other than investment in physical capital, which is included in some of the growth equations in Barro (1992)). Barro (1991) appeals to the work of Barro and Becker (1989) and Becker, Becker *et al.* (1990) to suggest that per capita growth and fertility move inversely. Other than this, no specific justification is given for the specification of the fertility equations.

[5] B–SM (1995: 453) argue that much of the reported movement over time in fertility, life expectancy and the infant mortality rate appears to be based on interpolation, so that changes in these variables over short periods of time (such as five- or ten-year spans) will contain little information. They, therefore, argue that it is more useful to consider observations that are relatively widely spaced in time, such as 1965 and 1985.

[6] In addition, B–L (1994) estimate a two-equation SURE system with fertility in 1985 and the change in fertility from 1965–1985 as the dependent variables. This paper focuses on the results obtained by jointly estimating cross-section equations for fertility in 1965 and 1985, as these are the results on which B–L place the most emphasis (and are the only results reported by B–SM).

[7] The total fertility rate measures 'the number of children that hypothetically would be born to a woman if she were to live to the end of her childbearing years and bear children at each age in accordance with the prevailing age-specific fertility rates' (World Bank, 1980: 14).

Table 1. *Re-estimation of the Barro–Lee fertility equation*

Variable	1	2	3	4
ln (GDP)	0.69**	0.815**	0.567*	0.091
	(2.76)	(3.08)	(2.09)	(1.71)
$(\ln(\text{GDP}))^2$	−0.053**	−0.061**	−0.043*	−0.013*
	(−3.12)	(−3.42)	(−2.32)	(−2.52)
FEMALE	−0.119**	−0.112**	−0.108*	−0.095*
	(−2.98)	(−2.60)	(−2.57)	(−2.20)
$(\text{FEMALE})^2$	0.0121**	0.011*	0.010*	0.009*
	(3.03)	(2.54)	(2.34)	(2.06)
MALE	0.155**	0.151**	0.118*	0.142**
	(3.52)	(3.13)	(2.44)	(2.92)
$(\text{MALE})^2$	−0.0145**	−0.014**	−0.010*	−0.013**
	(−3.45)	(−2.92)	(−2.20)	(−2.77)
ln (LE)	14.5*	13.3*	14.4*	14.3*
	(2.46)	(2.14)	(2.37)	(2.26)
$(\ln(\text{LE}))^2$	−1.88*	−1.75*	−1.89*	−1.86*
	(−2.47)	(−2.21)	(−2.44)	(−2.30)
MORT				
Both Years	7.2**	7.25**		7.43**
	(2.88)	(2.77)		(2.79)
1965			6.11*	
			(2.36)	
1985			11.5**	
			(3.85)	
$(\text{MORT})^2$				
Both Years	−25.7*	−27.4*		−28.0
	(−2.40)	(−2.38)		(−2.39)
1965			−20.8*	
			(−1.82)	
1985			−51.9**	
			(−3.33)	
R^2				
1965	0.81	0.81	0.82	0.80
1985	0.89	0.89	0.89	0.88
Tests of restrictions				
LR		19.52* (10)	9.87 (8)	23.31** (10)
Wald		16.62* (10)	8.19 (8)	20.57* (10)

Notes: dependent variable: log of the total fertility rate. SURE estimates with asymptotic t-statistics given in parentheses. **, * denote significance at the 1% and 5% levels respectively (against two-sided alternatives). Results in column 1 are from B–L (1994, Table 7, column 1) for their unbalanced panel. Results in columns 2 to 4 are for a balanced sample of 84 countries. Results in column 4 are for the B–L model estimated using Summers and Heston version 5.5 data. LR and Wald are the test statistics for the restrictions imposed in the relevant column; the figure in parentheses gives the degrees of freedom for each test.

i.e. the country-specific random error terms are correlated over time, but errors corresponding to different countries, either in the same year or different years, are uncorrelated.[8]

B–L jointly estimate Equations 1 and 2 and, without comment, restrict the coefficients on the explanatory variables to be equal across the two time periods, i.e. $\beta_{1k} = \beta_{2k}$, for $k = 1, \ldots, 10$.

Confirmation

B–L's sample consists of an unbalanced panel with 81 countries for 1965 and 89 countries for 1985. The results reported in B–L (1994, Table 7, column 1) are reproduced in column 1 of Table 1. B–L give details of the countries included in their growth equations, but not for their fertility equation.

[8] Note that B–L (and B–SM) do not explicitly address the potential endogeneity of the explanatory variables in their fertility equations. For example, if decreases in fertility promote growth in per capita income, estimated coefficients on per capita income are likely to be biased upwards.

In re-estimating their fertility equation, all countries for which the data were available for both years were included.[9] This gave a sample of 84 countries (listed in the Appendix); the potential loss of efficiency from excluding four countries for 1985 was traded off against the ease of computation with a balanced panel.[10] Results for the restricted model are reported in Table 1, column 2. These are qualitatively similar to B–L's original results. The coefficients on all the variables, including the non-linear terms, are statistically significant. At low levels of female schooling, an increase in female schooling reduces the fertility rate; however, the effect becomes weaker the higher the initial value of female schooling. A point is reached where an increase in female schooling leads, *ceteris paribus*, to a higher rate of fertility (the turning point in the original B–L results is at 4.9 years of female schooling). That the relationship can become positive at relatively low levels of schooling is a puzzling result, for which B–L offer no explanation. For male schooling, an increase in schooling leads to a higher fertility rate at low levels of schooling, but, at higher initial levels of schooling, an increase in male schooling leads to a reduction in the fertility rate (the turning point occurs at 5.3 years of male schooling).

The cross-equation restrictions imposed by B–L were tested using likelihood ratio (LR) and Wald tests. Under the null hypothesis, that the parameters on r variables in the two equations are equal, both the test statistics have an asymptotic χ^2 distribution with r degrees of freedom. The results are reported in Table 1. Restricting the coefficients on each of the explanatory variables to be equal for both years is rejected by both tests at the 5% significance level. Tests of each individual restriction (i.e. restricting the coefficients on only one variable at a time to be the same in both time periods) suggest that it is invalid to restrict the coefficients on both MORT and MORT^2.[11] The joint null hypothesis involving both these restrictions was also rejected on the basis of the LR and Wald tests. However, the null was not rejected on the basis of either test for any of the other variables. The results of SURE estimation with only the data-acceptable restrictions imposed are given in Table 1, column 3. All explanatory variables and their squared terms (except MORT^2 in 1965) are again significant, albeit at a lower level of significance in most cases.

Diagnostics

SURE estimation is more efficient than ordinary least squares (OLS) if there exists some correlation between the error terms for each country across time. The null hypothesis that there exists a diagonal covariance matrix was tested using the Breusch–Pagan (1980) Lagrange Multiplier (BPLM) test and an LR test. Each test has an asymptotic χ^2 distribution with one degree of freedom under the null. The test statistics for the model with data-acceptable restrictions (Table 1, column 3) are 27.147 for the BPLM test and 35.916 for the LR test. Therefore, both tests decisively reject the null hypothesis of a diagonal covariance matrix, so there is evidence of a correlation between the country-specific error terms in Equations 1 and 2. Estimating the equations separately, using OLS, will result in a loss of efficiency due to this and because the cross-equation restrictions are not being exploited. However, given the wider range of diagnostics available using OLS, the equations were also estimated separately.

The results for a selection of tests for heteroscedasticity, normality of the errors and model mis-specification for the separate equations for 1965 and 1985 are reported in Table 2. Each heteroscedasticity test has an asymptotic χ^2 distribution under the null hypothesis of homoscedasticity. For 1965, the null hypothesis of homoscedasticity is rejected by all the tests and for 1985 it is rejected by half of the tests. This may be indicative of a mis-specified model, as mis-specification often leads to evidence of heteroscedasticity (McAleer, 1994: 350). Normality of the error terms was tested using the Jarque–Bera (1980) Lagrange Multiplier normality test, which has a χ^2 distribution with two degrees of freedom under the null of normality. Normality of the errors is not rejected at the 5% significance level for either year. RESET tests for model mis-specification were also calculated. Each test has an F distribution under the null of correct model specification. The null is rejected for 1965 but not for 1985.

Lorgelly and Owen (1999) have shown that the significance of the male and female education variables in B–L's growth equations is sensitive to the sample of countries used. Tests for influential observations and/or outliers were, therefore, carried out for their fertility equations. On the basis of the studentized residuals and leverage, as measured by the diagonal values of the 'hat matrix' (Belsley *et al.*, 1980; Donald and Maddala, 1993), 20 countries were identified as being potentially influential observations and/or outliers. When the model corresponding to column 2 in Table 1 is re-estimated with these countries omitted, the coefficient on male schooling (and on MORT and MORT^2) becomes insignificant (although the coefficient on the squared term of male schooling remains signif-

[9] However, Kuwait was not included, which B–L (1994: 11) exclude from the sample for their growth equation because it 'was too unusual to include'; one assumes Kuwait is also excluded from the sample for their fertility equation.

[10] B–L do not provide details of how they computed the restricted SURE estimator for the unbalanced panel. SHAZAM version 7.0 was used for all computations in this paper.

[11] For $\beta_{19} = \beta_{29}$, LR = 4.529 and Wald = 3.526. For $\beta_{1,10} = \beta_{2,10}$, LR = 6.659 and Wald = 5.266. The critical value at the 5% significance level in each case is 3.841.

Table 2. *Diagnostic tests for the Barro–Lee and Barro and Sala-i-Martin fertility equations*

	B–L		B–SM	
Null hypothesis	1965	1985	1965	1985
Homoscedasticity				
$(\varepsilon_i)^2$ on $\hat{Y}_i$	7.62** (1)	0.00 (1)	2.50 (1)	0.23 (1)
$(\varepsilon_i)^2$ on $[\hat{Y}_i]^2$	9.30** (1)	0.18 (1)	3.37 (1)	1.43 (1)
$(\varepsilon_i)^2$ on $\ln[\hat{Y}_i]^2$	5.85* (1)	0.25 (1)	1.64 (1)	0.13 (1)
$(\varepsilon_i)^2$ on X	21.33* (10)	24.32** (10)	8.49 (6)	18.56** (6)
$\ln[(\varepsilon_i)^2]$ on X	18.83* (10)	36.04** (10)	6.28 (6)	8.05 (6)
$\lvert\varepsilon_i\rvert$ on X	22.16* (10)	28.93** (10)	11.76 (6)	18.32** (6)
Normal Errors				
Jarque Bera LM	0.68 (2)	0.07 (2)	8.53* (2)	0.67 (2)
Correct Specification				
RESET(2)	4.14* (1,72)	2.23 (1,72)	0.21 (1,76)	3.43 (1,76)
RESET(3)	7.15** (2,71)	1.88 (2,72)	8.49** (2,75)	2.32 (2,75)
RESET(4)	4.75** (3,70)	1.27 (3,70)	6.16** (3,74)	2.03 (3,74)

Notes: Degrees of freedom are given in parentheses. * and ** denote rejection of the specified null hypothesis at the 5% and 1% levels respectively. The heteroscedasticity and normality tests are χ^2 distributed and the RESET tests are F distributed under their respective nulls. $\hat{Y}_i$ represents the fitted value of ln (FERT) and X the relevant set of regressors.

icant).[12] With the influential observations omitted, the full set of cross-equation restrictions imposed by B–L are valid, and heteroscedasticity is no longer a problem in the 1985 equation. However, the 1965 equation still performs poorly in terms of diagnostics.[13] On the other hand, deleting observations that are influential for the coefficients on FEMALE and its squared value, i.e. those with high DFBETAS values (Belsley *et al.*, 1980), does not qualitatively affect the level of significance of the education variables compared to the full-sample results.

The B–L fertility equation has a reasonably high explanatory power and all the coefficients on the explanatory variables appear to be significant. However, the diagnostic test results suggest that there are problems with heteroscedasticity and model specification, and the significance of some of the variables is sensitive to the choice of sample. This suggests that any inferences drawn from the model, including tests of significance and the tests of cross-equation restrictions, should be treated with caution.

III. THE BARRO AND SALA-I-MARTIN MODEL

The fertility model that B–SM (1995) specify for 1965 and 1985 is given by:

$$\begin{aligned}\ln(\mathrm{FERT})_{1j} &= \alpha_{10} + \alpha_{11}\ln(\mathrm{GDP}_{1j}) + \alpha_{12}(\ln(\mathrm{GDP}_{1j}))^2 \\ &\quad + \alpha_{13}\mathrm{MPRIM}_{1j} + \alpha_{14}\mathrm{FPRIM}_{1j} \\ &\quad + \alpha_{15}\mathrm{MHIGH}_{1j} + \alpha_{16}\mathrm{FHIGH}_{1j} + \varepsilon_{1j} \qquad (3)\end{aligned}$$

$$\begin{aligned}\ln(\mathrm{FERT})_{2j} &= \alpha_{20} + \alpha_{21}\ln(\mathrm{GDP}_{2j}) + \alpha_{22}(\ln(\mathrm{GDP}_{2j}))^2 \\ &\quad + \alpha_{23}\mathrm{MPRIM}_{2j} + \alpha_{24}\mathrm{FPRIM}_{2j} \\ &\quad + \alpha_{25}\mathrm{MHIGH}_{2j} + \alpha_{26}\mathrm{FHIGH}_{2j} + \varepsilon_{2j} \qquad (4)\end{aligned}$$

MPRIM and FPRIM are, respectively, average years of male and female primary schooling, and MHIGH and FHIGH are, respectively, average years of male and female post-primary schooling. Definitions of the other variables are as for Equations 1 and 2.

This model differs in many respects from that of B–L. B–SM include a squared income term but do not include a squared term for any of the other explanatory variables in the model. B–L include total average years of schooling (for both males and females), whereas B–SM disaggregate average years of schooling into two categories (for both males and females): primary and post-primary (i.e. secondary and higher) education.[14] Life expectancy and infant mortality, both of which are found to be statistically significant by B–L, are omitted by B–SM; B–SM do not offer an explanation for any of the changes in the choice of

[12] The emphasis here is on assessing the robustness of the results to sample selection. It is not necessarily appropriate to omit such observations (see Lorgelly and Owen (1999) for further discussion).
[13] Details of the results are available on request.
[14] Disaggregating education does, however, allow for the possibility of a non-linear relationship. Coefficients of opposite sign on primary and post-primary education would be compatible with the type of non-linear relationship found by B–L.

Table 3. *Re-estimation of the Barro and Sala-i-Martin fertility equation*

Variable	1	2	3
ln (GDP)			
Both years	0.93**	0.993**	
1965	(3.00)	(3.07)	
			0.859*
1985			(2.17)
			0.325
			(0.85)
$(\ln(\mathrm{GDP}))^2$			
Both years	−0.070**	−0.080**	
	(−3.50)	(−3.76)	
1965			−0.067*
			(−2.55)
1985			−0.040
			(−1.62)
MPRIM	0.094**	0.076*	0.056
	(2.61)	(1.99)	(1.53)
FPRIM	−0.194**	−0.130**	−0.135**
	(−5.11)	(−3.19)	(−3.41)
MHIGH	−0.191**	−0.133*	−0.123*
	(−3.18)	(−2.06)	(−1.97)
FHIGH	0.155*	0.100	0.123
	(2.31)	(1.38)	(1.74)
R^2			
1965	0.68	0.69	0.72
1985	0.81	0.81	0.82
Tests of restrictions			
LR		24.85** (6)	9.64* (4)
Wald		23.15** (6)	8.82 (4)

Notes: Dependent variable: log of the total fertility rate. SURE estimates with asymptotic t-statistics given in parentheses. **, * denote significance at the 1% and 5% levels respectively (against two-sided alternatives). Results in column 1 are from B–SM (1995, Table 12.6, column 1) for their unbalanced panel. Results in columns 2 and 3 are for a balanced sample of 84 countries. See also notes to Table 1.

regressors or functional form compared to the B–L specification. Another difference is that the data on GDP per capita are from a later version of the Summers and Heston (1991) data set (version 5.5), whereas B–L use version 4 data.

Confirmation

B–SM estimate the model in Equations 3 and 4 as a SURE system using data on 90 countries for 1965 and 102 countries for 1985. The results reported by B–SM (1995, Table 12.6, column 1) are reproduced in column 1 of Table 3. B–SM do not give details of the countries included in the sample used to estimate their fertility equations. All countries for which the data are available for both years in the B–L data set (except for Kuwait) were included. The replicated results are based on a sample of 84 countries (listed in the Appendix) and are reported in column 2 of Table 3. The coefficient on female post-primary education is not statistically significant at the 5% level, whereas it is significant in B–SM.[15]

As for B–L, B–SM restrict the coefficients to be equal across the two time periods. B–SM report the results of testing the cross-equation restrictions on their growth equations, but not for their fertility equations. The results of testing the cross-equation restrictions for the sample are reported in Table 3. Both the LR and Wald tests reject the null hypothesis that the joint restrictions are valid. Tests of the restrictions on each pair of coefficients separately failed to isolate individual restrictions that are not data acceptable. However, jointly restricting the coefficients on both GDP and $(\mathrm{GDP})^2$ to be equal across time is rejected on the basis of both tests. The SURE results, for the model with the restrictions imposed for all the variables except GDP

[15] This is consistent with the sensitivity of results on the role of female education in the B–L growth equations to the sample coverage; see Lorgelly and Owen (1999).

and its squared term, are given in Table 3, column 3. Only the coefficients on income and its squared value for 1965, and female primary attainment and male post-primary attainment remain statistically significant. The original findings are not robust.

Diagnostics

The null hypothesis of a diagonal covariance matrix was tested using the BPLM test and an LR test (BPLM = 30.268 and LR = 42.327 for the model given in Table 3, column 3). These results suggest that to estimate the equations separately, using OLS, will result in a loss of efficiency. However, for ease of diagnostic testing, the equations were also estimated separately.

Heteroscedasticity, normality and RESET test results are reported in Table 2. The null hypothesis of homoscedasticity is not rejected for 1965, but is rejected by two of the tests for 1985. The RESET test results suggest that the 1965 equation may well be mis-specified. The LM normality test rejects the null hypothesis of normally distributed errors in 1965 but not for 1985.

On the basis of studentized residuals and leverage, 21 countries were identified as being influential observations or outliers. The model corresponding to Table 3, column 2 was re-estimated, with these observations omitted and only the data acceptable restrictions imposed (it is not valid to restrict the coefficients on $\ln(\mathrm{GDP})$, $(\ln(\mathrm{GDP}))^2$, MPRIM or FPRIM to be equal across time). The coefficients on MPRIM and FPRIM become statistically insignificant at the 5% level for 1965. When observations that are influential for FPRIM and FHIGH, on the basis of their DFBETAS values, are deleted, the results are similar; MPRIM and FPRIM remain statistically insignificant for 1965 (when the data-acceptable cross-equation restrictions are imposed).

The B–SM model does not appear to be particularly robust. While there appear to be fewer problems with heteroscedasticity, there is evidence of non-normality of the errors for 1965, which could affect the reliability of inference. Even if the unrestricted model is maintained, when only the (apparently) data-acceptable restrictions are imposed, several of the coefficients on the explanatory variables are no longer statistically significant. Again, the pattern of statistical significance is affected by the choice of sample. As with the B–L results, it appears that the B–SM results should also be treated with a good deal of caution.

IV. SPECIFICATION CHANGES: B–L VERSUS B–SM

B–SM (1995) do not explain why they choose to omit life expectancy and infant mortality from their fertility equation, given that these variables are found to be statistically significant by B–L. Some light could possibly be shed on this by estimating B–SM's equation, but with life expectancy and infant mortality included; the results are given in Table 4. In order to clarify B–SM's motivation for respecifying their model, B–L's and B–SM's preferred strategy was followed, i.e. imposing the full set of cross-equation restrictions. The results obtained when life expectancy and infant mortality (but not their squared values) are included are reported in column 1. The squared terms are included in the model in column 2.

In column 1, life expectancy is not statistically significant and infant mortality is only marginally significant (at the 5% level). In column 2, both life expectancy terms are significant, but both infant mortality terms are not significant. The non-significance of life expectancy and infant mortality may well be due to multicollinearity, as both are, essentially, proxies for the general state of health of the population. The simple correlation coefficient between the two variables is 0.90 for 1965 and 0.91 for 1985. That multicollinearity is a potential problem is confirmed in columns 3 and 4. When either life expectancy or infant mortality is dropped from the model, both the level and squared term of the remaining variable are highly significant. Therefore, it is not clear why B–SM chose to omit both the health variables from their fertility equations.

Diagnostic tests for the equations reported in Table 4 were also examined. Each equation suffers from potential problems from either heteroscedasticity (only the model in column 1 for 1965 has no significant heteroscedasticity test statistic values) or model mis-specification (all models for 1965, but only the model in column 1 for 1985). Non-normality of the errors is a problem for only the model in column 4 for 1965.

B–SM use a more recent version of the Summers and Heston income per capita data than B–L (version 5.5 compared to version 4). The results obtained from estimating the B–L model using version 5.5 data on output per capita are reported in Table 1, column 4. For all the variables other than income per capita, the size and statistical significance of the estimated coefficients are very similar to the comparable results in column 2. This is not altogether surprising as the sample and the data for the dependent variable and the explanatory variables, other than income per capita, are identical. However, while the coefficient on $(\ln(\mathrm{GDP}))^2$ is still statistically significant at the 5% level with a negative (though, in absolute terms, smaller) point estimate, the point estimate on the level term, $\ln(\mathrm{GDP})$, is drastically reduced and is not statistically significant at the 5% level (for a two-tailed test). This much more marginal result for income per capita, when using version 5.5 income data, may have been a relevant factor underlying B–SM's re-specification of their fertility equation; cf. B–SM's results in Table 3, column 1, where the coefficient on

Table 4. *Comparisons of the Barro–Lee and Barro and Sala-i-Martin specifications*

Variable	1	2	3	4
ln (GDP)	1.25**	0.858**	0.919**	0.744*
	(3.64)	(2.88)	(3.10)	(2.52)
$(\ln(\text{GDP}))^2$	−0.094**	−0.061**	−0.065**	−0.056**
	(−4.27)	(−3.15)	(−3.37)	(2.92)
MPRIM	0.097*	0.071*	0.058	0.099**
	(2.54)	(2.14)	(1.78)	(2.96)
FPRIM	−0.120**	−0.087*	−0.081*	−0.114**
	(−2.90)	(−2.45)	(−2.26)	(−3.19)
MHIGH	−0.111	−0.116*	−0.115*	−0.123*
	(−1.74)	(−2.10)	(−2.06)	(−2.24)
FHIGH	0.088	0.127*	0.125*	0.126*
	(1.23)	(2.06)	(2.01)	(2.05)
ln (LE)	0.062	18.3**	28.4**	
	(0.17)	(2.78)	(7.60)	
$(\ln(\text{LE}))^2$		−2.39**	−3.68**	
		(−2.86)	(−7.75)	
MORT	2.56*	5.08		12.3**
	(1.99)	(1.86)		(7.61)
$(\text{MORT})^2$		−22.4		−48.1**
		(−1.84)		(−6.73)
R^2				
1965	0.71	0.80	0.90	0.76
1985	0.84	0.88	0.88	0.88
Tests of restrictions				
LR	39.67** (8)	25.29** (10)	17.84** (8)	20.71* (8)
Wald	37.18** (8)	22.83* (10)	16.34* (8)	19.38* (8)

Notes: Dependent variable: log of the total fertility rate. SURE estimates with asymptotic t-statistics given in parentheses. **, * denote significance at the 1% and 5% levels respectively (against two-sided alternatives). All results are for a balanced panel of 84 countries. See also notes to Table 1.

ln (GDP) is significant at the 1% level (on a two-tailed test) and has a point estimate of 0.93.[16]

V. RE-ESTIMATING THE FERTILITY EQUATIONS USING UPDATED EDUCATION DATA

B–L (1996) have produced an updated cross-country data set on the educational attainment of the population. Both B–L (1994) and B–SM (1995) use B–L's (1993) earlier data on educational attainment. The new data set provides estimates of the educational attainment of those aged 15 and over; the earlier data relate to the population aged 25 and over. In both data sets, census data provide information on educational attainment for about 40 per cent of the data cells. In B–L (1993), missing cells are filled using information on gross enrolment rates. In B–L (1996), missing cells are filled using information on net enrolment rates. This latter approach avoids double counting due to students repeating grades or returning after previously dropping out (both of which lead to overestimation of the human capital stock).

B–L (1996: 218) emphasise that focusing on the population aged over 25, when for many developing countries a large proportion of the population is aged under 25, is a shortcoming of their previous data set. Therefore, it would seem more sensible to use the data for the population aged 15 and over in empirical studies of growth, fertility, etc.

B–L (1996: 222) note that they 'want especially to re-examine some previous puzzling results which failed to isolate a positive relation between female schooling and growth'. Somewhat surprisingly, Barro (1997), in the first study to use the new education data, finds that, in his growth equation, '[s]chooling of those aged twenty-five and over has somewhat more explanatory power than schooling of those aged fifteen and over' (Barro, 1997: 122, fn. 10).[17]

[16]Note that base-period values of GDP per capita are also not significant in four out of five fertility and population growth equations in Table II of Barro (1991). However, as noted in fn. 4, the explanatory variables in these equations are the same as those in the growth equations and are not likely to be optimal in explaining fertility.

[17]By contrast, in his regressions explaining an index of democracy, 'primary schooling of males and females aged fifteen and over has slightly more explanatory power than primary schooling of those aged twenty-five and over' (Barro, 1997: fn. 11).

Table 5. *Using alternative schooling data in the Barro–Lee model*

Variable	1	2
ln (GDP)	0.798**	0.784**
	(2.93)	(2.96))
(ln (GDP))^2	−0.059**	−0.059**
	(−3.28)	(−3.36)
FEMALE	−0.108*	−0.060
	(−2.49)	(−1.26)
(FEMALE)^2	0.011*	0.006
	(2.42)	(1.17)
MALE	0.108*	0.077
	(2.30)	(1.51)
(MALE)^2	−0.011*	−0.007
	(−2.44)	(−1.40)
ln (LE)	13.3*	15.3*
	(2.12)	(2.45)
(ln (LE))^2	−1.75*	−2.01*
	(−2.18)	(−2.52)
MORT	6.59*	6.37*
	(2.48)	(2.38)
(MORT)^2	−25.4*	−25.3*
	(−2.17)	(−2.15)
R^2		
1965	0.81	0.81
1985	0.88	0.88
Tests of restrictions		
LR	17.71 (10)	17.49 (10)
Wald	15.00 (10)	15.17 (10)

Notes: Dependent variable: log of the total fertility rate. SURE estimates with asymptotic *t*-statistics given in parentheses. **, * denote significance at the 1% and 5% levels respectively (against two-sided alternatives). Results in columns 1 and 2 are for the B–L model using B–L's (1996) 25+ and 15+ education data respectively. Results in column 1 are for a balanced panel of 83 countries, as data for Tanzania were not available. Results in column 2 are for the standard set of 84 countries. See also notes to Table 1.

B–L (1996) also state that they intend to use the new data to examine the effect of adult schooling on fertility. As noted by Schultz (1994: 257), 'one-half of childbearing occurs before women reach age 25, and as many as one-third of the women over 15 are between 15 and 25'. Therefore, using measures of schooling for those aged 25 and over may bias the effect of education on fertility, since a significant part of a woman's fertile life is excluded from the coverage of the education data. It is of interest to check how robust the fertility equations of B–L and B–SM are to the use of education data covering those aged 15 and over. Barro (1997) does not report estimates of a fertility equation, but refers (p. 21) to, amongst others, B–L (1994) for evidence on the role of female education in determining fertility.

Estimates for the B–L model, using the B–L (1996) education data are reported in Table 5. The benchmark results using the B–L (1993) data set are given in Table 1, column 2. The results in Table 5, column 1 are based on the new education data for the population aged 25 and over; the results in column 2 are based on the new data for those aged 15 and over. The full set of cross-equation restrictions appears to be data-acceptable for the models in columns 1 and 2; however, see also fn. 18. In column 1, all variables remain statistically significant (although, in some cases, at a lower level of significance) and of the same sign. While one would expect that attainment of females aged 15 and over would be a more relevant determinant of fertility, none of the 15 and over educational attainment variables is statistically significant in column 2. This is a puzzling result and casts doubt on either the validity of the B–L equation or of the 15 and over data, or, possibly, both.[18]

For the B–SM model, the benchmark results, using the B–L (1993) data, are given in Table 3, column 2. The results obtained from using the more recent education data for those aged 25 and over, and 15 and over, are given in Table 6, columns 1 and 2 respectively. In both cases, all the education variables, except female primary schooling, are no longer statistically significant.[19,20] Therefore, the B–SM equation appears to be even less robust than the B–L equation when re-estimated using the more recent education data.

VI. CONCLUSION

One of the aims of this paper was to establish why B–SM (1995) chose to omit variables from their fertility equation that are found to be significant in B–L (1994). One possible motivation identified is that income per capita is not statistically significant in the B–L fertility equation when Summers and Heston's version 5.5 cross-country income data are used instead of version 4 data. Given the widespread use of these data in empirical growth modelling, the

[18]For both the equations using 25+ and 15+ data, heteroscedasticity appears to be a problem in both years (particularly for 1965), and model mis-specification appears to be a problem in 1965. Diagnostics for the equations estimated using 25+ and 15+ data are available on request.

[19]Diagnostics for the equations estimated using 25+ and 15+ data suggest that, for both data sets, non-normal errors and model mis-specification are problems for 1965, and heteroscedasticity and model mis-specification are problems in 1985. Results are available on request.

[20]When only the (apparently) data-acceptable cross-equation restrictions are imposed on column (1) (LR and Wald test results suggest that it is invalid to restrict the coefficients on ln(GDP) and (ln(GDP))^2 to be constant over time), the coefficient on ln(GDP) is no longer statistically significant (for both time periods) and the coefficient on (ln(GDP))^2 is not significant for 1985. The same set of cross-equation restrictions was found to be valid for column 2. In this case, neither ln(GDP) nor (ln(GDP))^2 is significant for 1985. Thus, the results using the new data are even more disappointing when only the data-acceptable cross-equation restrictions are imposed.

Table 6. *Using alternative schooling data in the Barro and Sala-i-Martin model*

Variable	1	2
ln (GDP)	0.892**	0.991**
	(2.70)	(3.19)
$(\ln(\text{GDP}))^2$	−0.073**	−0.080**
	(−3.35)	(−3.94)
MPRIM	0.029	0.055
	(0.73)	(1.32)
FPRIM	−0.088*	−0.101*
	(−2.20)	(−2.37)
MHIGH	−0.057	−0.066
	(−0.86)	(−0.92)
FHIGH	0.005	0.007
	(0.06)	(0.10)
R^2		
1965	0.70	0.69
1985	0.80	0.80
Tests of restrictions		
LR	19.13** (6)	21.28** (6)
Wald	17.39** (6)	19.23** (6)

Notes: Dependent variable: log of the total fertility rate. SURE estimates with asymptotic *t*-statistics given in parentheses. **, * denote significance at the 1% and 5% levels respectively (against two-sided alternatives). Results in columns 1 and 2 are for the B–SM model using B–L's (1996) 25+ and 15+ education data respectively. Results in column 1 are for a balanced panel of 83 countries, as data for Tanzania were not available. Results in column 2 are for the standard set of 84 countries. See also notes to Table 1.

nature of the revisions between different vintages of the Summers-Heston data, and the effects on growth-equation estimates, may be worthy of further investigation. It is also shown that both the B–L and B–SM fertility equations impose apparently invalid cross-equation restrictions; any inferences should, however, be treated with caution as the estimated equations do not perform well in terms of diagnostic testing, which suggests mis-specification.

B–L (1996) strongly suggest that their education data are superior to the B–L (1993) data used in both B–L (1994) and B–SM (1995); however, when we use the new data, the results from both B–L and B–SM's fertility equations appear to be substantially weaker. In the B–L (1994) equation (which Barro (1997) cites as evidence of the role of female education in determining fertility), none of the education variables is significant in the determination of fertility when data for the population aged 15 and over are used. In B–SM's model, only female primary schooling is (negatively) correlated with fertility. The sensitivity of the different formulations of Barro's fertility equation to the use of different vintages of the educational attainment proxies and of the Summers–Heston cross-country income data is a cause for concern. It suggests that a robust explanation of fertility, to link with empirical growth equations, has not yet been found and that further work is required in this area.

ACKNOWLEDGEMENTS

The authors are grateful for financial support from the Marsden Fund (The Royal Society of New Zealand) and from a University of Otago Research Grant. This paper was originally circulated as University of Otago Economics Discussion Paper No. 9814. Paula Lorgelly is currently with the School of Economics and the Trent Institute for Health Services Research, School of Community Health Sciences, University of Nottingham.

REFERENCES

Barro, R. J. (1991) Economic growth in a cross section of countries, *Quarterly Journal of Economics*, **106**, 407–43.

Barro, R. J. (1992) Human capital and economic growth, in *Policies For Long-Run Economic Growth*, A Symposium Sponsored By The Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, pp. 199–216.

Barro, R. J. (1997) *Determinants of Economic Growth: A Cross-Country Empirical Study*, MIT Press, Cambridge, MA.

Barro, R. J. and Becker, G. S. (1989) Fertility choice in a model of economic growth, *Econometrica*, **57**, 481–501.

Barro, R. J. and Lee, J.-W. (1993) International comparisons of educational attainment, *Journal of Monetary Economics*, **32**, 363–94.

Barro, R. J. and Lee, J.-W. (1994) Sources of economic growth, *Carnegie-Rochester Conference Series on Public Policy*, **40**, 1–46.

Barro, R. J. and Lee, J. W. (1996) International measures of schooling years and schooling quality, *American Economic Review*, **86**, 218–23.

Barro, R. J. and Sala-i-Martin, X. (1995) *Economic Growth*, McGraw-Hill, New York.

Becker, G. S., Murphy, M. and Tamura, R. (1990) Human capital, fertility, and economic growth, *Journal of Political Economy*, **98**, S12–37.

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley, New York.

Breusch, T. S. and Pagan, A. R. (1980) The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies*, **47**, 239–53.

Caselli, F., Esquivel, G. and Lefort, F. (1996) Reopening the convergence debate: a new look at cross-country growth empirics, *Journal of Economic Growth*, **1**, 363–89.

Dewald, W. G., Thursby, T. G. and Anderson, R. G. (1986) Replication in empirical economics: the Journal of Money, Credit and Banking project, *American Economic Review*, **76**, 587–603.

Donald, S. G. and Maddala, G. S. (1993) Identifying outliers and influential observations in econometric models, in *Handbook of Statistics, Volume II* (Eds) G. S. Maddala, C. R. Rao and H. D. Vinod, Elsevier, Amsterdam, pp. 663–701.

Jarque, C. M. and Bera, A. K. (1980) Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters*, **6**, 255–9.

Lorgelly, P. K. and Owen, P. D. (1999) The effect of female and male schooling on economic growth in the Barro–Lee model, *Empirical Economics*, **24**, 537–57.

McAleer, M. (1994) Sherlock Holmes and the search for truth: a diagnostic tale, *Journal of Economic Surveys*, **8**, 317–70.

Schultz, T. P. (1994) Human capital, family planning, and their effects on population growth, *American Economic Review*, **84**, 255–60.
Subbarao, K. and Raney, L. (1995) Social gains from female education: a cross-national study, *Economic Development and Cultural Change*, **44**, 105–28.
Summers, R. and Heston, A. (1991) The Penn World Table (Mark 5): an expanded set of international comparisons, 1950–1988, *Quarterly Journal of Economics*, **106**, 327–68.
Tomek, W. G. (1993) Confirmation and replication in empirical econometrics: a step toward improved scholarship, *American Journal of Agricultural Economics*, **75**, 6–14.
World Bank (1980) *World Tables*, Johns Hopkins University Press, Baltimore.

APPENDIX

Data sources

The following equalities define the variables in terms of the variable names used in the B–L (1994) panel data set (available from http://www.nber.org/pub/barro.lee/). Detailed sources of their data are given in the file readme.txt as part of the data set. For $i = 1$, $xx = 65$ and for $i = 2$, $xx = 85$.

$$\ln(\mathrm{FERT}_i) = \ln(\mathrm{FERT}xx)$$
$$\ln(\mathrm{GDP}_i) = \ln(\mathrm{GDPSH4}xx)$$
$$\mathrm{FEMALE}_i = \mathrm{HUMANF}xx$$
$$\mathrm{FPRIM}_i = \mathrm{PYRF}xx$$
$$\mathrm{FHIGH}_i = \mathrm{SYRF}xx + \mathrm{HYRF}xx$$
$$\mathrm{MALE}_i = \mathrm{HUMANM}xx$$
$$\mathrm{MPRIM}_i = \mathrm{PYRM}xx$$
$$\mathrm{MHIGH}_i = \mathrm{SYRM}xx + \mathrm{HYRM}xx$$
$$\ln(\mathrm{LE}_i) = \ln(\mathrm{LIFEE0}xx)$$
$$\mathrm{MORT}_i = \mathrm{MORT}xx$$

In the B–SM replication, the income variable is from version 5.5 of the Summers and Heston data set and is defined in the B–L panel data set as:

$$\ln(\mathrm{GDP}_i) = \ln(\mathrm{GDPSH5}xx)$$

The following equalities define the variables in terms of the variable names used in the 1996 B–L education data set (available from http://www.worldbank.org/research/growth/ddbarle2.htm). Detailed sources of their data are given in the readme file as part of the data set. The data on the education of those aged 25 and over are given in the school2.raw and school3.raw files, using the observations for which YEAR = 65 and 85.

$$\mathrm{FEMALE}_i = \mathrm{TYRF25}$$
$$\mathrm{FPRIM}_i = \mathrm{PYRF25}$$
$$\mathrm{FHIGH}_i = \mathrm{SYRF25} + \mathrm{HYRF25}$$
$$\mathrm{MALE}_i = \mathrm{TYRM25}$$
$$\mathrm{MPRIM}_i = \mathrm{PYRM25}$$
$$\mathrm{MHIGH}_i = \mathrm{SYRM25} + \mathrm{HYRM25}$$

The data on the education of those aged 15 and over are given in the school5.raw and school6.raw files, where YEAR = 65 and 85.

$$\mathrm{FEMALE}_i = \mathrm{TYRF15}$$
$$\mathrm{FPRIM}_i = \mathrm{PYRF15}$$
$$\mathrm{FHIGH}_i = \mathrm{SYRF15} + \mathrm{HYRF15}$$
$$\mathrm{MALE}_i = \mathrm{TYRM15}$$
$$\mathrm{MPRIM}_i = \mathrm{PYRM15}$$
$$\mathrm{MHIGH}_i = \mathrm{SYRM15} + \mathrm{HYRM15}$$

Countries included: Algeria, Botswana, Ghana, Kenya, Lesotho, Liberia, Malawi, Mauritius, Mozambique, Niger, Senegal, Sierra Leone, South Africa, Sudan*, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zaire, Zambia, Zimbabwe, Canada, Costa Rica, Dominican Republic, El Salvador, Guatemala, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Trinidad and Tobago, the USA, Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Uruguay, Venezuela, Bangladesh, Hong Kong, India, Indonesia, Israel, Japan, Jordan, Korea, Malaysia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Syria, Thailand, Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the UK, Yugoslavia**, Australia, Fiji, New Zealand, Papua New Guinea

** excluded from the B–L replication, * excluded from the B–SM replication.