Oxford Economic Papers 54 (2002), 118–149

Are educational gender gaps a brake on economic development? Some cross-country empirical evidence

By Stephen Knowles*, Paula K. Lorgelly*, and P. Dorian Owen†

* Department of Economics, University of Otago
† Department of Economics, University of Otago, PO Box 56, Dunedin, New Zealand; email: dowen@business.otago.ac.nz

This paper estimates a neoclassical growth model that includes female and male education as separate explanatory variables. The model can be reparameterised so that the gender gap in education enters the model. The interpretation of its coefficient depends crucially on what other education variables appear in the equation. The average long-run effects of female and male education on output per worker are estimated for a cross section of countries using long time averages of the data. The results support the World Bank's emphasis on the importance of female education in raising labour productivity and are robust to various sensitivity analyses.

1. Introduction

> If we educate a boy, we educate one person. If we educate a girl, we educate a family—and a whole nation. (African proverb, quoted by James Wolfensohn, President of the World Bank, 1995).

It has almost become part of the conventional wisdom of development economics that, in developing countries, the economic gains from educating females are greater than from educating males. The essence of this view is captured in the above African proverb quoted by Wolfensohn (1995) and is reflected in the World Bank's strategies and policies towards gender and development (World Bank, 1994). Wolfensohn goes on to argue that educating girls has a 'catalytic effect' on every dimension of economic development, including higher productivity and faster economic growth. However, in many developing countries, especially in Africa and South Asia, males enjoy considerably higher levels of schooling than females; i.e. there is an educational gender gap (see Appendix 2).[1] If the World

[1] Note that the data presented in Appendix 2 are averages for the period 1960–1990 and are for the mean years of schooling of the whole population aged 15 and over. To the extent that school enrolments have increased over time, the mean years of schooling of school leavers in 1990 will be greater (in many countries considerably greater) than the figures reported in the Appendix.

Bank view is correct, then virtually all developing countries are not investing in the optimal mix of female and male education.[2]

The World Bank view is consistent with educational rate-of-return analyses using micro data. For example, Psacharopoulos (1994) finds that the rate of return to female education is positive, and marginally higher than that to male education. However, rate-of-return analyses using micro data on individual earnings are likely to understate the true contribution of female schooling to labour productivity if there are benefits from female education that have indirect effects on measured output that are not captured in females' earnings.

There are good reasons to believe that female and male education affect output levels and growth in different ways. Female education, as with male education, can improve productivity when better-educated females participate in the paid workforce; this contributes directly to conventionally measured output and, indirectly, through the flow-on effects of higher output to higher physical capital investment. Female participation rates are generally lower than for males and vary widely across countries, but are likely to increase with higher levels of educational attainment. However, there is evidence that female education, especially in developing countries, also produces social gains by reducing fertility and infant mortality, improving family and child health, increasing life expectancy, and increasing the quantity and quality of children's educational attainment (e.g., Schultz, 1988; Behrman and Deolalikar, 1988; Subbarao and Raney, 1995). Hence, even if female participation rates are lower than for males, the effects of improved female education on general levels of education, health status, and population growth can boost measured productivity growth indirectly.[3]

Studies using cross-country aggregate macro-level data should, in theory, pick up not only the direct effects of female education on conventionally measured labour productivity, but also the indirect effects not captured using micro data.[4] However, the cross-country macroeconomic empirical literature seems far from reaching a consensus on the role of female education. For example, Hill and King (1993, 1995) present evidence that the level of female education has a significant positive effect on GNP and also that larger gender gaps in school enrolments reduce GNP. By contrast, in one of the most cited sequences of recent empirical growth

[2] The reasons why gender gaps in education persist are reviewed by, for example, Hill and King (1995), Schultz (1995), and Dollar and Gatti (1999).

[3] For given levels of (female or male) attainment, gender inequality in educational attainment could itself have negative effects on output and growth, e.g., by lowering the average ability levels of those educated compared to a situation of equality. Such inequality has been interpreted as acting like a distortionary tax that leads to a misallocation of educational resources (Dollar and Gatti, 1999) and has flow-on effects in terms of access to employment and technology (Klasen, 1999). Gender inequality in education could also reduce the potential for synergies in educational attainment between siblings or couples (Klasen, 1999).

[4] Any empirical measure of the total effects on measured productivity growth, even using aggregate data, would ignore any positive effects of female education on unmeasured productivity in home production. Indeed, it should be noted that increases in measured productivity arising from improved female education may partly reflect a substitution from unmeasured to measured production.

studies (Barro and Lee, 1994; Barro and Sala-i-Martin, 1995; Barro, 1996a, b), Barro and his colleagues find that the partial correlation between female secondary schooling and economic growth is negative, whereas for male schooling the partial correlation is positive; this appears to be the opposite to Hill and King's conclusions. However, a major problem with interpreting the results of these studies, and the other existing cross-country studies discussed below, is that the equations estimated are *ad hoc*; see, for example, Sims' (1996) comments on Barro (1996a). It is not clear how the coefficients on female education, male education, and/or an educational gender gap are to be interpreted.

The aim of this paper is to obtain empirical estimates of the different long-run effects of female and male schooling on labour productivity within the context of a model with a more explicit theoretical structure than existing studies. Male and female education are included in a neoclassical growth model in the tradition of Solow (1956), Swan (1956), and Mankiw, Romer, and Weil (1992) (hereafter MRW). The long-run effects of the education variables are defined in terms of aggregate output elasticities with respect to female and male schooling and physical capital, which have an obvious interpretation. We also provide an explicit interpretation of coefficients on educational gender gaps.

Existing empirical results on the effects of gender-separate education variables on measured output or growth are reviewed in Section 2. The theoretical model is outlined in Section 3. The estimation method employed and the data are discussed in Section 4, and the basic results are reported in Section 5. In Sections 6 and 7, we compare our results with the existing literature and present the results of a range of sensitivity analyses, including examining the effects of influential observations and simultaneity. Section 8 concludes.

2. Review of existing empirical work

The most obvious feature of the results from existing cross-country macroeconomic studies is their lack of agreement on the effects of gender-specific education on growth and/or output levels.[5] Some studies report OLS regressions implying positive effects of both female and male education, e.g., Benavot (1989) for female and male primary school enrolments on growth in *per capita* GNP, although the coefficients on both female and male secondary schooling are generally not statistically significant. Point estimates suggest that female primary enrolments have larger effects on growth than male enrolments, especially for the African, Latin American, and poor LDC sub-samples. However, standard errors are not reported and none of the cross-section regressions includes both female and male enrolments in the same equation. Schultz (1995, p. 20), in some largely

[5] Some studies estimate growth equations and others levels equations. The results from both are relevant because, as Hall and Jones (1999, p. 114) note '[m]any of the predictions of growth theory can be successfully considered in a cross-section context by examining the levels of income across countries'. See also footnote 15.

descriptive regressions that do not control for other conventional growth determinants, also finds that female school enrolments have an apparently greater positive impact on economic growth than male school enrolments. Birdsall *et al.* (1997), in a variant of Barro's (1991) cross-section growth regression, find that the coefficients on female and male primary enrolments are both positive, jointly (though not individually) significant, and not significantly different from each other.

Hill and King (1993, 1995) explicitly introduce an educational gender gap into a cross-section output regression. They regress income *per capita* on the gender gap, female secondary school enrolment, and a set of control variables.[6] Their results suggest that the level of female education has a significant positive impact on the level of GNP and that 'large gender disparities in educational attainment are associated with lower levels of GNP' (Hill and King, 1995, p. 29). That is, for given levels of female education, higher levels of male education reduce GNP. A more recent study by Klasen (1999) also includes proxies for both the general level of education and the gender gap in a cross-country growth equation. He finds significant positive coefficients on the female/male ratio of base-period average years of schooling and the female/male ratio of growth in years of schooling over a 32-year period; his estimated equations also include base-period years of schooling for males, and the growth in years of schooling for males over the period.

By contrast, Barro and Lee (1994) find that there is a significant negative partial correlation between economic growth and base-period female secondary schooling and a significant positive partial correlation between economic growth and male secondary schooling. They use Barro and Lee's (1993) average years of schooling data, rather than enrolment rates, and seemingly unrelated regression equations and instrumental variables estimation methods applied to cross-country data for two time periods (1965–75 and 1975–85). They describe their negative coefficient on female schooling as a 'puzzling finding' but suggest that 'a high spread between male and female secondary attainment is a good measure of backwardness; hence, less female attainment signifies more backwardness and accordingly higher growth potential through the convergence mechanism' (Barro and Lee, 1994, p. 18).[7] This negative partial correlation is representative of the results from a number of other studies by Barro (Barro and Sala-i-Martin, 1995; Barro, 1996a,b) that use different definitions of the years of schooling variables

[6] They define the gender gap as 'the ratio of female to male enrolment at the primary or secondary level, whichever is smaller (that is, where the gap is larger)' (Hill and King, 1995, p.14). Rather than measuring the gender gap continuously, Hill and King include it in the regression equation using three (intercept) dummy variables (defined by whether the gender gap is in the ranges <0.42, 0.42–0.75, 0.75–0.95).

[7] This explanation is not entirely convincing given that they also include base-period income *per capita* in their regressions to model the convergence mechanism. An alternative explanation, suggested by Stokey (1994) and further examined by Lorgelly and Owen (1999), is that their result is largely due to the influence of four Asian countries (Hong Kong, Singapore, Taiwan, and Korea) that had very high levels of growth but very low initial levels of female schooling.

(secondary, higher, secondary + higher, etc.) and include additional control variables in essentially the same framework. A similar pattern of results is reported by Perotti (1996) who estimates a cross-section regression including the distribution of income as an additional explanatory variable.

This group of studies by Barro and his colleagues includes life expectancy to proxy for the health component of human capital. When education is not disaggregated by gender, Knowles and Owen (1995, 1997) find that education is not statistically significant in a range of models that include life expectancy. Barro (1996b, p. 6) confirms this result, noting that '[i]f a measure of health status, such as life expectancy, is included in the regression, then it seems to proxy for the level of human capital. The level of educational attainment then has no additional explanatory power for growth. An additional positive effect on growth emerges, however, when male attainment is high relative to female attainment'. On this interpretation, Barro reaches completely the opposite conclusion, compared to Hill and King (1993, 1995) and Klasen (1999), on the effect of the educational gender gap on growth.[8]

In a further twist, female education is not included as an explanatory variable in the preferred results reported by Barro (1997, 1998) who notes that, when the more recent Barro and Lee (1996) education data are used, 'the estimated female coefficients are essentially zero' (Barro, 1997, p. 122).[9]

In response to problems involved in estimating cross-section growth equations, Caselli *et al.* (1996) re-estimate Barro and Lee's equations using a generalised method of moments (GMM) estimator, the pros and cons of which are discussed in Section 4. They use panel data including five five-year periods (whereas Barro and Lee use two ten-year periods). Caselli *et al.* obtain a statistically significant positive coefficient on female schooling and a significant negative coefficient on male schooling, i.e. the opposite of Barro and Lee's result. Their argument (p. 379) is that 'the female education variable captures both (positive) fertility effects, and (negative) human capital effects, and the former outweighs the latter. Male

[8] Dollar and Gatti (1999) also include male and female educational attainment separately with life expectancy in a growth equation estimated using panel data. For the full sample, the coefficients on the male and female educational attainment variables are the opposite of those suggested by Barro, but are not statistically significant. When the sample is split in half, on the basis of female educational attainment, female attainment has a positive statistically significant effect on growth but only for the 'more developed' half of the sample, for which the coefficient on male attainment is negative but not statistically significant. This bifurcation may partly reflect Dollar and Gatti's choice of educational proxy (the percentage of the population for whom secondary education is the highest level attained). Surprisingly, their data appendix suggests that they used the 1993 vintage of Barro and Lee's educational data for the population over the age of 25.

[9] The 1996 data set provides estimates of the educational attainment of those aged 15 and over; the 1993 data relate only to the population aged 25 and over. In both data sets, census data provide information on educational attainment for about 40% of the data cells. Missing cells are filled using information on gross enrolment rates in Barro and Lee (1993) and on net enrolment rates in Barro and Lee (1996). 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).

education, on the other hand, only represents a human capital effect. Hence, its negative coefficient'.[10] Forbes (2000), in an examination of the effect of inequality on growth, applies similar estimation methods to Caselli *et al.*, but uses Barro and Lee's (1996) years of schooling data; she obtains a statistically significant positive effect of female schooling and a negative, but not significant, effect of male schooling.

Other studies, e.g., Knight *et al.* (1993) using five-year panels, have also found (gender-aggregated) education to have a significant negative effect. Knight *et al.* argue that this is due to five-year time intervals being too short to pick up the effect of school enrolment rates on growth (as it will be some time before those currently enrolled in secondary school enter the workforce). They argue (p. 533) that 'cross-sectional data (data where the observation for each country is an average of its respective time-series observations) may be the preferred proxy in estimating the growth effects of human capital investment'. Islam (1995) finds a marginally significant and negative correlation between average years of schooling and economic growth. He suggests that one possible reason for this is that years of schooling is a better proxy for human capital when used in a pure cross section, rather than where there is a time-series dimension. Temple (1999b, p. 133) also argues that 'the effect of human capital has to be almost immediate if it is to be detected in a panel'. Business cycle effects, which are typically ignored when panel data are used, are also likely to be relevant when using data with a frequency as short as five years.

Part of the reason for the different results may, therefore, be whether a short-run or long-run effect of education on growth (or output) is being measured, which will be influenced by the frequency of the data and by whether the estimation method involves differencing the data. Our focus of interest is the long-run effect of education on productivity and our approach is based on the view that cross-country variation in the levels of the variables is important in assessing the strength of this relationship.

Although existing aggregate cross-country studies on the effects of female education on growth and output levels are contradictory, one thing they unfortunately all have in common is a lack of an explicit theoretical framework. This makes the coefficients on the schooling variables, particularly educational gender gaps, difficult to interpret. In the next section we include gender-separate education in an augmented neoclassical model. While not without its problems, the extended neoclassical model has the advantage of being a well-used and well-understood framework.

[10] While this recognises the different indirect effects of female education on growth, the characterisation of the 'human capital effect' as negative runs counter to most arguments on the sign of this effect (see, for example, Topel, 1999; Krueger and Lindahl, 1998); only the convergence-proxy argument suggested by Barro and Lee (1994), or a negative exogenous change in the returns to schooling, are consistent with a negative effect.

3. The theoretical framework

We include gender-specific education in a MRW-style neoclassical growth model. Following Barro and Lee (1994) and Knowles and Owen (1995, 1997) we broaden the definition of human capital to include health as well as education. The aggregate production function is given by

$$Y_{it} = K_{it}^{\alpha} EF_{it}^{\beta_f} EM_{it}^{\beta_m} X_{it}^{\psi} (A_{it} L_{it})^{1-\alpha-\beta_f-\beta_m-\psi} \tag{1}$$

where Y is real output, K the stock of real physical capital, EF the stock of female education, EM the stock of male education, X the stock of health capital, A the level of technology, and L the labour force.[11] Subscripts i and t denote country i and time period t respectively. The production function is Cobb-Douglas, exhibits constant returns to scale, and the marginal products of each factor are assumed to be positive and diminishing.[12] As will be discussed in Section 4, it is likely that the elasticities, α, β_f, β_m, and ψ, vary across different countries; however, for notational simplicity we suppress the country subscripts on the parameters.

Equation (1) can be written as

$$y_{it} = k_{it}^{\alpha} ef_{it}^{\beta_f} em_{it}^{\beta_m} x_{it}^{\psi} \tag{2}$$

where lower case letters denote quantities per effective unit of labour (e.g., $y = Y/AL$). Following MRW, the labour force and technology are assumed to be determined by

$$L_{it} = L_{i0} e^{n_i t} \tag{3}$$

$$A_{it} = A_{i0} e^{gt} \tag{4}$$

where n is the growth rate of the labour force and g the growth rate of technology (assumed to be constant across countries). The accumulation of physical capital, female and male education, and health are given by

$$\dot{k}_{it} = s_{ki} y_{it} - (n_i + g + \delta) k_{it} \tag{5}$$

$$\dot{ef}_{it} = s_{efi} y_{it} - (n_i + g + \delta) ef_{it} \tag{6}$$

$$\dot{em}_{it} = s_{emi} y_{it} - (n_i + g + \delta) em_{it} \tag{7}$$

$$\dot{x}_{it} = s_{xi} y_{it} - (n_i + g + \delta) x_{it} \tag{8}$$

where s_{ki}, s_{efi}, s_{emi}, and s_{xi} are the fractions of real output invested in physical capital, female education, male education, and health respectively. δ is the rate of

[11] Here male and female education, and health capital are entered as separate factors in the production function. Alternative functional forms are possible. Knowles and Owen (1997), for example, estimate a model in which education and health are labour-augmenting, but their results are qualitatively similar to a separate-factors specification. Another alternative is to use a Mincerian formulation (e.g., Bils and Klenow, 1998; Hall and Jones, 1999). This specification is discussed in more detail in Section 6.2.

[12] The assumption of constant returns to scale was tested by estimating a more general version of eq. (2), which did not impose constant returns to scale, for which the data were available for 47 countries. (For many of the countries in the sample, data were not available for the capital stock per worker.) The null hypothesis of constant returns to scale was not rejected.

depreciation, which is assumed to be common across countries and time. Assuming the existence of a steady state (with $\alpha + \beta_f + \beta_m + \psi < 1$), eqs. (5) to (8) imply

$$k_i^* = \left(\frac{s_{ki}^{1-\beta_f-\beta_m-\psi} s_{efi}^{\beta_f} s_{emi}^{\beta_m} s_{xi}^{\psi}}{n_i + g + \delta} \right)^{1/\eta} \tag{9}$$

$$ef_i^* = \left(\frac{s_{ki}^{\alpha} s_{efi}^{1-\alpha-\beta_m-\psi} s_{emi}^{\beta_m} s_{xi}^{\psi}}{n_i + g + \delta} \right)^{1/\eta} \tag{10}$$

$$em_i^* = \left(\frac{s_{ki}^{\alpha} s_{efi}^{\beta_f} s_{emi}^{1-\alpha-\beta_f-\psi} s_{xi}^{\psi}}{n_i + g + \delta} \right)^{1/\eta} \tag{11}$$

$$x_i^* = \left(\frac{s_{ki}^{\alpha} s_{efi}^{\beta_f} s_{emi}^{\beta_m} s_{xi}^{1-\alpha-\beta_f-\beta_m}}{n_i + g + \delta} \right)^{1/\eta} \tag{12}$$

where an asterisk denotes steady-state values and $\eta \equiv 1 - \alpha - \beta_f - \beta_m - \psi$. Substituting eqs. (4) and (9) to (12) into eq. (2), taking natural logarithms and rearranging gives

$$\ln\left(\frac{Y_{it}}{L_{it}}\right)^* = \ln A_{i0} + gt - \frac{1-\eta}{\eta}\ln(n_i + g + \delta) + \frac{\alpha}{\eta}\ln(s_{ki}) + \frac{\beta_f}{\eta}\ln(s_{efi}) + \frac{\beta_m}{\eta}\ln(s_{emi}) + \frac{\psi}{\eta}\ln(s_{xi}) \tag{13}$$

Equation (13) expresses steady-state output per worker[13] as a function of the savings rate for each factor of production. As will be discussed in Section 4, for estimation purposes it is preferable to have an equation where the human capital variables enter as steady-state stocks. Solving eqs. (10) to (12) for their respective savings rates, as functions of the steady-state stocks, and substituting these into eq. (13) gives

$$\ln\left(\frac{Y_{it}}{L_{it}}\right)^* = \ln A_{i0} + gt - \frac{\alpha}{1-\alpha}\ln(n_i + g + \delta) + \frac{\alpha}{1-\alpha}\ln(s_{ki}) + \frac{\beta_f}{1-\alpha}\ln(ef_{it}^*) + \frac{\beta_m}{1-\alpha}\ln(em_{it}^*) + \frac{\psi}{1-\alpha}\ln(x_{it}^*) \tag{14}$$

Subsuming the growth rate of technology into a constant (a) and an error term (ε_i) gives the following equation for estimation[14]

[13] $(Y/L)^*$ denotes values of output per worker compatible with steady-state output per effective worker. Given the assumptions made, 'steady-state' output per worker grows at rate g.

[14] Commonly, base-period technology is also subsumed into the constant and the error term. We attempt to avoid the serious estimation problems this causes by incorporating a proxy for the variation in $\ln A$ across countries.

$$\ln\left(\frac{Y_{it}}{L_{it}}\right)^{*} = a + \ln A_{i0} + \frac{\alpha}{1-\alpha}(\ln(s_{ki}) - \ln(n_i + g + \delta)) + \frac{\beta_f}{1-\alpha}\ln(ef_{it}^{*}) + \frac{\beta_m}{1-\alpha}\ln(em_{it}^{*}) + \frac{\psi}{1-\alpha}\ln(x_{it}^{*}) + \varepsilon_{it} \tag{15}$$

Note that the coefficients on $\ln(s_{ki})$ and $\ln(n_i + g + \delta)$ sum to zero. A test of this restriction provides some guidance as to the adequacy of the model. If the null hypothesis of a valid restriction is accepted, then the restricted version of the equation can be used to infer values of α, β_f, β_m, and ψ.

Barro and Lee (1994), Barro and Sala-i-Martin (1995), and Barro (1996a,b) include female and male years of schooling as separate explanatory variables, similar to eq. (15). Hill and King (1993, 1995) include female schooling and the gender gap. Our model can be reparameterised so that education enters in a similar fashion to Hill and King

$$\ln\left(\frac{Y_{it}}{L_{it}}\right)^{*} = a + \ln A_{i0} + \frac{\alpha}{1-\alpha}(\ln(s_{ki}) - \ln(n_i + g + \delta)) + \frac{\beta_f + \beta_m}{1-\alpha}\ln(ef_{it}^{*}) + \frac{\beta_m}{1-\alpha}(\ln(em_{it}^{*}) - \ln(ef_{it}^{*})) + \frac{\psi}{1-\alpha}\ln(x_{it}^{*}) + \varepsilon_{i} \tag{16}$$

If $0 < \alpha < 1$ and $0 < \beta_m < 1$, the expected sign of the coefficient on the gender gap is positive. However, the interpretation of this coefficient is different from that implicit in Hill and King's discussion. In the context of our model, when the stock of female education is also included in the equation, a highly significant coefficient on the gap term does not imply anything about the role of female education *per se*, rather it reflects the output elasticity with respect to male education, as well as α. The output elasticity with respect to female education has to be unscrambled from the coefficient on the stock of female education in which it is combined with β_m and α. A negative coefficient on the gender gap requires either $\beta_m < 0$ or $\alpha > 1$.

Alternatively, the model can be reparameterised so that education enters as male schooling and a gender gap, in a similar fashion to Klasen (1999, Tables 2 and 4)

$$\ln\left(\frac{Y_{it}}{L_{it}}\right)^{*} = a + \ln A_{i0} + \frac{\alpha}{1-\alpha}(\ln(s_{ki}) - \ln(n_i + g + \delta)) + \frac{\beta_f + \beta_m}{1-\alpha}\ln(em_{it}^{*}) - \frac{\beta_f}{1-\alpha}(\ln(em_{it}^{*}) - \ln(ef_{it}^{*})) + \frac{\psi}{1-\alpha}\ln(x_{it}^{*}) + \varepsilon_{it} \tag{17}$$

For this parameterisation, if $0 < \alpha < 1$ and $0 < \beta_f < 1$, the expected sign of the coefficient on the gender gap is negative. The coefficient on the gender gap now depends on β_f and α, but not β_m. The quote from Barro (1996b) in Section 2 suggests that the levels of male and/or female educational attainment have no additional explanatory power, but the gap between female and male attainment has an effect; this requires $\beta_f = -\beta_m$ in eq. (16) or (17). A key message of eqs. (16) and (17) is that the expected sign of the coefficient on the gender gap depends crucially on what other education variables are included in the estimated equation.

4. Estimation issues and data

4.1 Estimation

We interpret the steady-state relationship in eq. (14) as an 'attractor' or long-run equilibrium relationship between the variables. Even though this is the central equation in the neoclassical growth model, it is specified in terms of levels of the variables and contains the parameters of interest, i.e. the elasticities on the human capital components (β_f, β_m, ψ).[15] A key issue is how best to estimate the parameters in this long-run levels relationship.

Partly because of the over-riding interest in convergence, much of the existing empirical growth literature involves estimation of dynamic 'Barro regressions', after Barro (1991), in which base-period output per worker (or *per capita*) is included in cross-section (or 'small T' panel) regressions. Most of the studies discussed in Section 2 are of this type. However, it is has become clear that this approach has serious conceptual and econometric problems (Pesaran and Smith, 1995; Caselli *et al*., 1996; Sims, 1996; Lee *et al*., 1997). In cross-section estimation any unobservable country-specific effects (including cross-country differences in technology, if not proxied explicitly) enter the error terms in the estimated equations. OLS (and GLS-type methods such as SURE estimation) yield consistent estimates only if the country-specific error terms are uncorrelated with the explanatory variables. In conventional dynamic cross-section regressions the errors are necessarily correlated with base-period output *per capita* (Caselli *et al*., 1996, p. 367) which can severely bias estimates of both the convergence parameter and the underlying production function elasticities.[16] Because of these problems, we consider it undesirable to rely on extracting estimates for the elasticities in the steady-state relationship from the coefficients estimated from a conventional

[15] In the standard neoclassical growth model, growth occurs only through exogenous technological progress or through the transitional dynamics of the convergence mechanism. The human capital variables affect growth through their effects on the steady-state level of output per worker. For example, in eq. (13), assuming β_f and η are positive, an increase in investment in female education increases steady-state output per worker and, for a given value of existing (or base-period) output per worker, this increases the transitional growth rate. Hence, these elasticities are crucial to both the growth and levels effects of changes in human capital.

[16] An additional source of bias arises from the potential endogeneity of the explanatory variables, as investment rates, population growth, schooling levels and life expectancy are all likely to depend on a country's level of output per worker. With the unobservable country-specific component of the level of technology, A, in the error term, estimated coefficients on these variables are likely to be biased upwards. A common response to endogeneity concerns is to use lagged values as instruments. However, this may not solve the problem as it is not obvious that they are necessarily uncorrelated with the country-specific (rather than a time-period-specific) error term. Indeed, Sims (1996, p. 176) emphasises that 'since the components of the error term determined by A are probably persistent, there are no grounds for concluding that these variables are actually predetermined'. For these reasons, we avoid use of lagged values when examining the sensitivity of our results to potential simultaneity and also include a proxy for A in most of the reported regressions.

cross-section growth equation which includes transitional dynamics (and hence initial output per worker).

A recent response to these problems is the use of GMM estimation of a dynamic panel specification (e.g., Caselli *et al.*, 1996; Forbes, 2000). However, the GMM estimators used in these studies have several drawbacks given that we are specifically interested in the long-run steady-state relationship between education and output per worker. The most important is that, while they remove the country-specific effects by differencing the time-series data, this also removes information about the long-run 'levels' relationship in the time dimension (Hendry, 1995, p. 287) and eliminates all the between-country variation in levels in the cross section.[17] The latter is of particular concern from the viewpoint that cross-section variation is informative about the long-run relationship between the variables (Baltagi and Griffin, 1984), especially as cross-section variation in the levels of the variables is generally greater than time-series variation given the relatively short time spans of data currently available.[18]

If annual time-series data were available for all the relevant variables a natural approach would be to use panel cointegration techniques, as the levels variables are likely to be non-stationary. Unfortunately, the data for female and male education are not available on an annual basis (and many of the observations for life expectancy are interpolated). Instead, we use cross-section regressions with time-averaged data to estimate the parameters in the long-run steady-state levels relationship. This is motivated by recent developments in regression limit theory for non-stationary panel data by Pesaran and Smith (1995), Moon and Phillips (1998), and Phillips and Moon (1999, 2000).

A feature emphasised in this literature is potential parameter heterogeneity across groups, which is likely in a cross-country context. Pesaran and Smith (1995) focus on the average long-run effects across countries; in the current context these represent the cross-country means of the functions of the elasticities in the steady-state relationships, assuming that these constitute valid cointegrating relationships. Where there is parameter heterogeneity, aggregation and pooling of the data in dynamic panels yield estimates of such average long-run effects that are inconsistent and potentially very misleading (Pesaran and Smith, 1995). However, Pesaran and Smith show that, if the coefficients vary randomly across countries and

---

[17] Differenced models 'entail the counter-intuitive prediction that agents do not try to remove disequilibrium in levels' (Hendry, 1995, p. 248), which, in the growth context, is not consistent with the importance of the gap between steady-state and initial levels of *per capita* output for transitional dynamics.

[18] Differencing also reduces the 'signal-to-noise' ratio in the data, which can exacerbate measurement error problems, especially for variables that change relatively slowly over time (Barro, 1996a; Temple, 1999a). In addition, the GMM estimator used by Caselli *et al.* (1996) and Forbes (2000) uses lagged values of the relevant stock and flow variables as instruments. While these are more likely to be uncorrelated with the error term in the differenced model, which includes components of ΔA rather than A, they are likely to be weak instruments for differences of these variables, again especially for variables that are highly persistent over time (Temple, 1999a).

are independent of the regressors, cross-section regressions based on long time averages can produce consistent estimates of the average long-run coefficients, although their results are based on restrictive assumptions, including strict exogeneity of the regressors.[19]

Phillips and Moon (1999, 2000) provide a different interpretation of long-run average relationships when using cross-section regressions of time averages in non-stationary panels with heterogeneous parameters. The panel regression coefficients are analogous to classical regression coefficients but the relations are parameterised in terms of the matrix regression coefficient of the long-run average covariance matrix for the cross section rather than, as in conventional regressions, the covariance matrix for the data. A key difference compared to Pesaran and Smith is that Phillips and Moon's long-run average regression coefficients are defined irrespective of whether cointegrating relationships exist in the time series for the individual countries or whether cointegrating relationships differ across countries. By contrast, for the average long-run effects examined by Pesaran and Smith such cointegrating time-series relations need to exist. Moon and Phillips (1998) show, under much weaker conditions than Pesaran and Smith, that limiting cross-section estimation using time-averaged data provides consistent estimates for the parameters in the long-run average relationship and that the estimator has a limiting normal distribution.[20] Our approach is, therefore, to estimate the static steady-state equation using long time averages of the data, and, following Phillips and Moon, to interpret the estimated coefficients as average cross-country long-run effects. As Temple (1999a, p. 126) notes '[g]iven that the purpose of cross-country empirical work is often to arrive at generalisations about growth, the averages are important'.

4.2 Data

Previous research has used data on school enrolment rates to proxy for investment in educational human capital (e.g., MRW, 1992), whereas data on average years of schooling have been widely used to proxy for stocks of educational human capital. Because school enrolments are not a particularly good proxy for accumulation of schooling (Gemmell, 1996; Pritchett, 1996) we estimate the variant of the model where the human capital variables enter as steady-state stocks rather than as savings rates. The education data used are Barro and Lee's (1996) data on the average years of schooling of the population aged 15 and over, disaggregated by gender. Their

[19] Estimating cross-section regressions purely from levels means that the transitional dynamics found in most empirical applications of the neoclassical growth model are ignored. However, Pesaran and Smith show that omitting such terms does not affect the consistency of the estimates of the average long-run coefficients.

[20] Phillips and Moon's analysis assumes cross-section independence. Further work on the implications for the asymptotic theory of cross-section dependence, for example arising from global or regional shocks, and evidence on finite-sample properties is clearly desirable.

data are available at five-yearly intervals from 1960 to 1990. Averages for the period 1960 to 1990, are reported in Appendix 2.[21,22]

Life expectancy has become a reasonably common proxy for health status in the literature on economic growth (e.g., Wheeler, 1980; Barro and Lee, 1994; Barro and Sala-i-Martin, 1995; Knowles and Owen, 1995, 1997). The life expectancy data (1965–90) are from the World Bank's *Social Indicators of Development*. Following Anand and Ravallion (1993) and Knowles and Owen (1995, 1997), the life expectancy data are transformed to take account of non-linearities in the data using the formula $\ln(x) = -\ln(85 - LE)$ where LE equals life expectancy at birth. Data on life expectancy measure mortality, rather than morbidity. It could be argued that when examining the relationship between health and labour productivity, data on morbidity would be preferable, but unfortunately these are not available for a large cross section of countries.[23] However, for a cross section of countries at different levels of economic development, morbidity and mortality are likely to be highly correlated.

Data on income per worker and the share of physical capital investment in national income are from the Penn World Tables (version 5.6). The growth rate of the labour force is calculated using the Penn World Tables data on income per worker, income *per capita* and population. Following MRW (1992) and Knowles and Owen (1995, 1997), $(g + \delta)$ is assumed to equal 5%. To proxy technical efficiency we use Hall and Jones's (1999) estimates derived as residuals from a levels accounting exercise. Any attempt to measure unobservable technical efficiency is fraught with difficulties but we contend that some explicit proxy for cross-country variation in $\ln A$ is better than none.[24] More details on data sources are given in Appendix 1.

5. Empirical results

The results obtained from OLS estimation of variants of eq. (15) using time averages over the period 1960–1990 are reported in Table 1. The asymptotic

[21] Data for each available year were first logged then averaged over the appropriate period. For some variables and countries, data were not available for all years, and averages over shorter periods were taken; however, countries with a sizeable proportion of data missing were excluded and care was also taken that variables with missing data were not biased toward earlier or later years.

[22] Although Barro and Lee's years of schooling data are widely used as proxies for educational human capital, recent work by de la Fuente and Doménech (2000), re-examining schooling data for a sample of OECD countries, suggests that they may be subject to measurement error.

[23] Sen (1998) argues the data on morbidity that are available (which rely on people's perception of their own morbidity) are not consistent across countries because of the subjectivity involved. He suggests (p. 20) that '[w]e may get a much better idea of people's ability to avoid death and severe illness by looking at actual mortality information, rather than from self-perception of ailments'.

[24] There is some evidence to suggest that, for common samples, the correlations between different constructed proxies are reasonably high (de la Fuente and Doménech, 2000).

Table 1 OLS estimates of the steady-state equation for GDP per worker
Dependent variable: average log of income per worker 1960–1990 (except column (v))

Variable	(i)	(ii)	(iii)	(iv)	(v)
Unrestricted model					
$\ln(s_k)$	0.291†	0.089	0.219*	0.262**	0.401**
	(1.86)	(0.58)	(2.29)	(2.88)	(9.39)
$\ln(n+g+\delta)$	−1.213**	0.053	−0.149	0.054	−0.320**
	(−2.90)	(0.13)	(0.55)	(0.16)	(−2.78)
$\ln(ef)^*$	0.663**	0.321	0.443**	0.438**	0.219**
	(2.92)	(1.54)	(3.68)	(3.90)	(3.38)
$\ln(em)^*$	−0.048	−0.161	−0.236	−0.324†	0.079
	(−0.15)	(−0.54)	(−1.28)	(−1.71)	(1.00)
$\ln(x)^*$		1.292**	0.309†	0.256	0.166**
		(5.49)	(1.90)	(1.52)	(2.58)
$\ln(A)$			0.710**	0.744**	0.738**
			(9.86)	(10.05)	(20.96)
$\ln(Y_0/L_0)$					−0.810**
					(−25.16)
R^2 adjusted	0.696	0.782	0.909	0.862	0.935
Restricted model					
$\ln(s_k) - \ln(n+g+\delta)$	0.393**	0.081	0.215*	0.258**	0.395**
	(2.83)	(0.55)	(2.33)	(2.88)	(10.62)
$\ln(ef)^*$	0.625**	0.335	0.449**	0.469**	0.230**
	(2.69)	(1.62)	(3.80)	(4.14)	(3.74)
$\ln(em)^*$	0.001	−0.166	−0.238	−0.347†	0.076
	(0.004)	(−0.55)	(−1.29)	(−1.82)	(0.99)
$\ln(x)^*$		1.261**	0.294†	0.194	0.155*
		(6.00)	(1.83)	(1.17)	(2.41)
$\ln(A)$			0.710**	0.747**	0.737**
			(9.78)	(9.89)	(20.83)
$\ln(Y_0/L_0)$					−0.811**
					(−24.56)
R^2 adjusted	0.686	0.785	0.910	0.864	0.936
F	3.545†	0.106	0.062	0.889	0.374
N	73	72	72	51	72
Implied elasticities					
Implied α	0.282**	0.075	0.177**	0.205**	0.328**
	(3.95)	(0.59)	(2.83)	(3.62)	(16.50)
Implied β_f	0.449*	0.310	0.370**	0.373**	0.191**
	(2.54)	(1.54)	(3.60)	(3.89)	(3.91)
Implied β_m	0.001	−0.153	−0.196	−0.275†	0.063
	(0.004)	(−0.55)	(−1.28)	(−1.79)	(0.99)
Implied ψ		1.167**	0.242†	0.154	0.129*
		(3.83)	(1.68)	(1.12)	(2.28)

Notes: OLS estimates with heteroscedasticity-consistent t-statistics given in parentheses. Intercept terms are not reported. **, *, and † denote significance at the 1%, 5% and 10% levels respectively (against two-sided alternatives). F is the test statistic obtained for the null hypothesis that the restriction implied by the model is valid. N is the sample size. For the elasticities, asymptotic Wald t-statistics for the null hypothesis that the relevant elasticity equals zero are given in parentheses. The dependent variable in column (v) is $\ln(Y_{90}/L_{90}) - \ln(Y_{60}/L_{60})$.

t-statistics, reported in parentheses, are based on a multivariate extension of the asymptotic standard errors obtained in the limit theory for time-averaged regressions derived by Moon and Phillips (1998).[25] The model passes the test of the restriction that the coefficients on $\ln(s_{ki})$ and $\ln(n_i + g + \delta)$ sum to zero, apart from the variant in column (i) at the 10% significance level. The values of the implied estimates for the elasticities are also reported, with, in parentheses, a Wald test (calculated as an asymptotic t-test) for the null hypothesis that the relevant elasticity is zero.[26] Column (i) reports the results obtained when life expectancy and technology are excluded from the model. The coefficient on female education, and the implied β_f, are positive and statistically significant, whereas the coefficient on male schooling, and β_m, are not significantly different from zero. The implication is that countries with higher levels of female schooling will have higher levels of labour productivity, *ceteris paribus*. The inclusion of life expectancy (column (ii)) renders the coefficients on all variables except life expectancy not statistically significant. The point estimate of ψ implied by the restricted regression is greater than unity.[27]

However, given the discussion in Section 4.1, omission of a proxy for the level of technology, which is likely to vary across countries, is a potential source of bias in the results in columns (i) and (ii). In particular, it may well be that, in column (ii), life expectancy is partly proxying for unmeasured country-specific effects. To attempt to control for such bias, we include Hall and Jones's estimates of $\ln(A)$ in subsequent regressions. The fit of the equation, in column (iii), is improved, the coefficient on life expectancy is now significant only at the 10% level and, while the implied point estimate of ψ now takes on a much more plausible value, it is not statistically significant. Both α and β_f are significant at the 1% level.

Previous empirical work on economic growth (e.g. Temple, 1998a,b; Knowles and Owen, 1995) has found that coefficients are not always stable when the sample is split into high-income and less developed country (LDC) samples. Given that the

[25] The t-values are effectively equivalent to heteroscedasticity-consistent t-statistics, which Pesaran and Smith (1995, p. 94) also argue are appropriate for regressions with long time averages. Conventional t-values generally give qualitatively similar results to those reported, as is the case in Moon and Phillips' (1998) examination of cross-country savings-investment relationships.

[26] The Wald tests involve non-linear combinations of the coefficients in the estimated model and are not invariant to reparameterisation of the restrictions being tested. If θ is the parameter on the $[\ln(s_{ki}) - \ln(n_i + g + \delta)]$ term, then, since $\theta \equiv \alpha/(1-\alpha)$, a test of $\alpha = 0$ is equivalent to a test of $\theta = 0$, provided $\alpha \neq 1$ (e.g., see Knowles and Owen, 1997, p. 322). Hence the significance of the implied elasticities can also be inferred from the reported t-statistics on the appropriate estimated coefficient in the restricted model. While asymptotically equivalent they give different results in finite samples. However, we find they generally give qualitatively similar results.

[27] While this appears to be inconsistent with the assumption that there are diminishing returns to each factor of production, this is not the case if life expectancy (*LE*) is only a proxy for health capital (measured in real terms in units of foregone consumption). Hence if $LE = X^{\gamma}$, where $\gamma < 1$, then the coefficient on life expectancy in the production function could overstate the underlying 'true' elasticity of output with respect to health capital. We are grateful to Steve Dowrick for pointing this out.

World Bank's arguments regarding the role of female education are made in the context of developing countries, the results for the LDC sample are of particular interest. Column (iv) presents the results when the high-income countries (as classified by the World Bank, 1993) are omitted from the sample. The female education elasticity remains significant at the 1% level. The male education elasticity is again negative but significant only at the 10% level, and the elasticity for life expectancy is not statistically significant.[28]

The results reported so far are for equations explaining steady-state levels of output per worker estimated using long time averages. As noted in Section 4.1, an alternative approach is to extract estimates of the elasticities from an equation explaining the growth rate of output per worker in the transition to the steady state. Following MRW's (1992) method of linearising the model around the steady state, a dynamic equation corresponding to equation (14) is given by

$$\begin{aligned}\ln\left(\frac{Y_{it}}{L_{it}}\right) - \ln\left(\frac{Y_{i0}}{L_{i0}}\right) = {}& \theta \ln A_0 + gt + \frac{\theta\alpha}{1-\alpha}(\ln(s_{ki}) - \ln(n_i + g + \delta)) \\ & + \frac{\theta\beta_f}{1-\alpha}\ln(ef_{it}^*) + \frac{\theta\beta_m}{1-\alpha}\ln(em_{it}^*) \\ & + \frac{\theta\psi}{1-\alpha}\ln(x_{it}^*) + \theta\ln\left(\frac{Y_{i0}}{L_{i0}}\right) \qquad (18)\end{aligned}$$

where $\theta = (1 - \exp(-\lambda_i t))$ and λ is interpreted as a speed of convergence parameter (e.g., see Knowles and Owen, 1995). As discussed earlier, there are compelling arguments against estimating growth equations that contain base-period output per worker. However, for comparison purposes, we present estimates of eq. (18) in column (v) of Table 1. Despite reservations about this approach, the main results are broadly similar to those in column (iii). The main difference is that the implied point estimate of β_f is somewhat lower whereas the point estimate of α is higher in the dynamic equation.

The results obtained from estimating the reparameterised versions of the model in Table 1, column (iii) are given in Table 2, together with 95% confidence intervals for the implied elasticities, obtained using the method outlined in Knowles and Owen (1997, p. 323). Being alternative parameterisations these are statistically equivalent, so that there is no additional empirical content in these equations compared to those in Table 2, column (iii). However, they illustrate the need to take into account what other education variables are included in the estimated equation when interpreting the estimated coefficient on the educational gender gap. The lack of statistical significance of the coefficient on the gap in column (i) reflects the lack of statistical significance of β_m, whereas the statistical significance of the gap coefficient in column (ii) reflects the statistical significance of α and β_f. The results also highlight the implications of different degrees of

[28] Removing life expectancy from the equations estimated in (iii) and (iv) does not qualitatively affect the results, although the point estimates for α and β_f increase.

Table 2 OLS estimates of reparameterisations of the steady-state equation for GDP per worker
Dependent variable: average log of income per worker 1960–1990

	(i) Eq. (16)		(ii) Eq. (17)
Unrestricted model			
$\ln(s_k)$	0.219*		0.219*
	(2.29)		(2.29)
$\ln(n+g+\delta)$	−0.149		−0.149
	(−0.55)		(−0.55)
$\ln(ef)^*$	0.207†		
	(1.90)		
$\ln(em)^*$			0.207†
			(1.90)
$\ln(em)^* - \ln(ef)^*$	−0.236		−0.443**
	(−1.28)		(−3.68)
$\ln(x)^*$	0.309†		0.309†
	(1.90)		(1.90)
$\ln(A)$	0.710**		0.710**
	(9.86)		(9.86)
Restricted model			
$\ln(s_k) - \ln(n+g+\delta)$	0.215*		0.215*
	(2.33)		(2.33)
$\ln(ef)^*$	0.211†		
	(1.93)		
$\ln(em)^*$			0.211†
			(1.93)
$\ln(em)^* - \ln(ef)^*$	−0.238		−0.449**
	(−1.29)		(−3.80)
$\ln(x)^*$	0.294†		0.294†
	(1.83)		(1.83)
$\ln(A)$	0.710**		0.710**
	(9.78)		(9.78)
$CI(\alpha, 95)$		0.03, 0.28	
$CI(\beta_f, 95)$		0.18, 0.58	
$CI(\beta_m, 95)$		−0.50, 0.11	
$CI(\psi, 95)$		−0.02, 0.56	

Notes: CI(., 95) is the 95% confidence interval for the relevant elasticity. See also notes to Table 1.

multicollinearity for the statistical significance of the different combinations of levels/gap variables included. For the full sample, the simple correlation between male and female schooling is 0.948, compared to −0.901 between female schooling and the gender gap, and −0.717 between male schooling and the gender gap. The potential for multicollinearity to inflate standard errors is likely to be lower for the parameterisation where male schooling and the gender gap are the explanatory variables, and this is reflected in column (ii), where both male schooling (at the 10% level) and the gender gap (at the 1% level) are statistically significant in both the restricted and unrestricted equations. Note that the restriction $\beta_f = -\beta_m$, corresponding to coefficients equal to zero on ef^* in eq. (16) and on em^* in eq. (17), is not rejected (marginally) at the 5% level. This is consistent with a significant long-run effect of the educational gender gap on output per worker, but no additional effects from either of the separate levels of schooling.

Our empirical results, so far, appear to be consistent with the World Bank view that, on average across a wide range of countries, female education makes a greater contribution to labour productivity than male education.

6. Reconciling our results with the existing empirical literature

6.1 Reconciliation with Barro-style regressions

The implications for the role of female and male education differ substantially from those of Barro and his colleagues (Barro and Lee, 1994; Barro and Sala-i-Martin, 1995; Barro 1996a,b, 1997, 1998). It is, therefore, worth summarising the main differences between eq. (15) and a typical Barro-style regression. We do not include the same range of control variables as Barro but we do include a proxy for $\ln(A)$, which does not appear explicitly in Barro's growth regressions. We also use the more recent Barro and Lee (1996) education data, whereas most of Barro's earlier work uses the older Barro and Lee (1993) data. Barro (1997, 1998) uses the more recent data but finds that the effect of female schooling is not statistically significant. In addition, we average the data over the whole estimation period, whereas in Barro's work the explanatory variables are measured either at a point in time (usually the base year) or as an average over a ten year period. Finally, we focus on estimating a long-run steady-state equation for output per worker, whereas Barro estimates growth equations including proxies for transitional dynamics.

To explore the consequences of using different vintages of the education data we re-estimated eq. (15) using the Barro and Lee (1993) education data for the population aged 25 and over.[29] Life expectancy and $\ln(A)$ are both included in the model for all the estimates in Table 3, so the relevant comparison is with column (iii) of Table 1. The implied elasticities, reported in row (a) of Table 3, are similar: β_f remains positive and statistically significant, whereas β_m is not statistically significant. Using different vintages of the Barro and Lee education data does not appear to be the reason for the difference in results.

Next, we examined the effect of including additional control variables as Barro's work typically includes a larger set of control variables than we do. The rationale for including these variables is not always clear. For example, Barro and Lee (1994, p.19) argue that the number of revolutions 'influences property rights and thereby affects the incentive to invest in various activities'. As the rate of investment is already included as an explanatory variable, it seems unnecessary to also include the

[29] Given that many of those aged between 15 and 24 are likely to be in the workforce, especially in developing countries, we consider the data on average years of schooling of those aged 15 and over to be a preferable proxy to that of those aged 25 and over.

Table 3 Implied elasticities obtained from OLS estimation of the steady-state equation when performing various sensitivity analyses
Dependent variable: average log of income per worker 1960–1990

Sensitivity test	α	β_f	β_m	ψ
(a) using 1993 education data	0.222**	0.343**	−0.189	0.121
	(5.03)	(3.89)	(−1.39)	(1.00)
(b) adding extra control variables	0.280**	0.304**	−0.109	0.213
	(5.45)	(3.66)	(−0.83)	(1.49)
(c) non-averaged (1990) data	0.284**	0.276**	−0.00004	0.135*
	(11.06)	(5.27)	(−0.001)	(2.00)
(d) base-period human capital (levels equation)	0.345**	−0.006	0.162**	0.152**
	(18.28)	(−0.20)	(3.83)	(3.05)
(e) base-period human capital (growth equation)	0.405**	−0.073**	0.248**	0.059
	(24.52)	(−2.84)	(6.54)	(1.09)
(f) *RSTUDENT* and leverage	0.219**	0.409**	−0.175	0.108
	(3.64)	(4.38)	(−1.17)	(0.76)
(g) *DFBETAS* (female & male)	0.236**	0.256**	−0.028	0.164
	(6.14)	(4.29)	(−0.29)	(1.55)
(h) omitting NICs	0.176**	0.308**	−0.069	0.217
	(2.84)	(3.13)	(−0.49)	(1.60)
(i) LAE	0.239**	0.269**	−0.079	0.140
	(4.88)	(2.89)	(−0.58)	(1.08)
(j) LTS	0.137*	0.484**	−0.172	0.139
	(2.24)	(4.59)	(−1.18)	(1.01)

Notes: Asymptotic Wald t-statistics for the null hypothesis that the relevant elasticity equals zero are given in parentheses. For all rows except (f) these are heteroscedasticity-consistent. See also the notes to Table 1.

number of revolutions.[30] However, for comparability, we re-estimated eq. (15) and included the following set of control variables included in Barro-style regressions: government consumption (net of spending on education and defence) as a share of GDP, the number of revolutions per year, and the black market premium. The implied values of the elasticities, reported in row (b) of Table 3, are again qualitatively similar to those from column (iii) of Table 1.

Averaging the data over a long time span may account for some of the difference between our results and Barro's. The results obtained when eq. (15) is estimated including period $t(= 1990)$ values for output per worker and the human capital variables, rather than averages over the period, are reported in row (c).[31] This

[30] Consistent with the explicit use of an aggregate production function, we include factor inputs and technology as explanatory variables. Other variables reflecting preferences, institutional differences, etc., affect steady-state GDP per worker through their influences on factor inputs and technology (Jones, 1997, p. 136).

[31] To be consistent with, for example, MRW (1992) we use data for the terminal-period to proxy the steady-state values but continue to average the data on labour force growth and investment in physical capital over the period 1960–90.

change has relatively little effect except that ψ is significant at the 5% level and the implied point estimate for α is larger and β_f smaller. By contrast, Barro's studies include base-period values of 'state variables' including human capital stocks. Re-estimating the equations in Table 1, columns (iii) and (v) including base-period values (1960 for the educational variables and 1965 for life expectancy) gives the results reported in Table 3, rows (d) and (e). This change has a much more dramatic effect. For the levels equation, β_m becomes positive and statistically significant, whereas β_f is not statistically significant. In the growth equation, the results are similar, except that the negative β_f is statistically significant.[32] It appears that a major contributing factor in explaining the difference between our results and Barro's is his use of base-period values for the human capital stocks in his growth equations. The model in Section 3, however, implies that the timing of the starred steady-state values should be treated consistently, which is incompatible with using time t values for (Y/L) and time 0 values for the human capital stocks in eq. (15) and hence eq. (18). The difference in these results is also consistent with Stokey's (1994) explanation for Barro and Lee's (1994) negative effect of female education on growth. For the fast-growing East Asian countries, the gender gaps in education declined rapidly over the period, so that the use of base-period measures of educational human capital accentuates the extent of educational gaps for these countries; base-period female education ends up acting as a 'chopsticks dummy' (see also Lorgelly and Owen, 1999, p. 551; Klasen, 1999, p. 20).[33]

6.2 Reconciliation with Mincerian regressions

Some recent papers (e.g. Topel, 1999; Krueger and Lindahl, 1998; Klenow and Rodriguez-Clare, 1997) have suggested that the coefficients on human capital obtained in the macro growth literature should be treated with caution as they are unrealistically higher than those typically obtained from Mincerian regressions (using micro data). It is, therefore, important to consider whether our results are consistent with the labour literature. As the coefficients on male education tend not to be statistically significant in our results, we focus the discussion on female education, for which the estimates are more precise.

[32] Similar results to those in rows (d) and (e) are obtained if $\ln(A)$ is omitted, so the inclusion of a proxy for the state of technology is not crucial in reversing the Barro-Lee results; the timing of the human capital stocks is more important.

[33] Klasen (1999), despite including base-period values of both male years of schooling and the ratio of female to male years of schooling in his regressions finds, like us, that educational inequality impedes growth. The difference in results, compared to Barro's studies, is probably due to Klasen's inclusion of regional dummy variables in his regressions. Dollar and Gatti (1999, p. 19) also suggest that inclusion of regional dummies overturns Barro and Lee's (1994) finding of a negative coefficient on female education, although they argue that it is mainly due to poor growth and high levels of female secondary schooling in Latin America.

Mincerian regressions, following Mincer (1974), involve estimating the following equation

$$\ln W_{it} = a + \beta S_{it} + \varepsilon_{it} \tag{19}$$

where W_{it} and S_{it} are, respectively, the wage paid to and the years of schooling of individual i at time t.[34] If eq. (19) were to be estimated using macro data, then years of schooling for individual i would be replaced by average years of schooling for country i.

Mincerian regressions typically give estimates of the rate of return to education between 0.05 and 0.15 (Bils and Klenow, 1998; Topel, 1999).[35] These figures are considerably lower than the values of β_f obtained in this paper (typically around 0.3 to 0.4). However, we would not expect β_f to be identical to the estimates of β obtained in Mincerian regressions for, at least, two reasons. The first is that we expect our β_f to exceed estimates obtained from Mincerian regressions using micro data because of the positive externalities due to female schooling. Ignoring other factors, similar estimated returns from the two approaches require either that the effects of the externalities are not economically significant and/or that Mincerian regressions pick up the social as well as the private returns.

The second distinction is due to the different functional forms employed in eqs. (2) and (19). The log-log form of our regression equation (which follows directly from including average years of schooling as a factor of production in a Cobb-Douglas production function) means that β_f can be interpreted as an elasticity. This functional form allows for diminishing returns to extra years of schooling, consistent with the finding of Psacharopoulos (1994) that Mincerian returns tend to be lower in countries with higher average years of schooling. The semi-log functional form employed in a Mincerian regression (and typically used in Barro-style regressions) means that β measures the percentage increase in income due to a one unit (one year) increase in average years of schooling. When applied to a sample covering a wide range of levels of development, this is not consistent with diminishing returns to schooling. Hence, our estimates of β_f (typically around 0.3 to 0.4) are broadly consistent with the values in the region of 0.05–0.15 obtained in Mincerian regressions, either with or without a contribution from the indirect benefits from female education.[36]

[34] Some Mincerian regressions also control for the number of years an individual has been working to allow for the human-capital-enhancing effect of on-the-job training, as opposed to human capital attributed to the formal education system.

[35] See Temple (2000) for a discussion of the issues involved in interpreting Mincerian regressions and a comparison of the evidence on the role of education from Mincerian regressions and growth equations.

[36] For example, if average years of schooling increases from four to five years (the mean female years of schooling for our data sample is 4.37) this represents a 25% increase. A β_f of 0.35 implies an increase in income per worker of 8.75%. For this example, a β_f of 0.35 from a macro regression is consistent with a Mincerian return of 0.0875. Taking the upper and lower limits of the 95% confidence interval for β_f in Table 2 (0.18, 0.58) gives Mincerian returns in the range (0.045, 0.145). Our estimates of the productivity effects of female education therefore appear to be at least as large as those obtained from Mincerian regressions.

7. Sensitivity analyses

Previous work by Lorgelly and Owen (1999) has shown that Barro and Lee's (1994) empirical results are not robust if the East Asian newly industrialising countries are omitted from the data sample. To test the sensitivity of our results we re-estimated eq. (15) omitting observations that could be classified, using different criteria, as influential or outliers. The estimates of the implied elasticities from the reduced samples are reported in Table 3. The criteria used to select observations for deletion are: large studentized residuals (*RSTUDENT*) (>2) and/or high leverage ($>2k/N$, where k is the number of regressors and N the sample size) in row (f) (Belsley *et al.*, 1980); high values ($> 2/N^{1/2}$) for the coefficient-specific *DFBETAS* values (Belsley *et al.*, 1980) for female and/or male schooling in row (g); the three Asian NICs (Taiwan is not included in our full sample) in row (h). An alternative to identification of outliers and influential observations is accommodation by the use of robust estimation methods. Row (i) reports least absolute error (LAE) estimates (see, for example, Judge *et al.*, 1988) and row (j) least trimmed squares (LTS) estimates (Rousseeuw and Leroy, 1987).[37] While there is some variation in point estimates, the coefficients and the elasticities on female education are positive and significant (at the 5% level or better) while the coefficients on male education are not statistically significant for all these sets of estimates.

Simultaneity is another potential source of bias in coefficient estimates from cross-country regressions. While inclusion of a proxy for $\ln(A)$ reduces the likely correlation between the explanatory variables and omitted country-specific effects in the error term, it is quite possible that all of our explanatory variables are partly determined by income per worker, in which case they will be correlated with the country-specific error term. In particular, Pritchett and Summers (1996) argue, based on comparisons of sets of instrumental variables' estimates, that high cross-country correlations between *per capita* income and proxies for health status are due to increases in income raising health status.[38] The strong correlation between life expectancy and output per worker that dominates the results in Table 1, column (ii) are, however, significantly attenuated when the proxy for $\ln(A)$ is included in the regression model. Whereas most of the literature emphasizes the effect of schooling on growth, Bils and Klenow (1998) argue that the correlations between growth and education variables arise from causality running from growth to schooling.

In order to assess the sensitivity of our results to simultaneity bias, we re-estimated the model using conventional two-stage least squares (2SLS). For this

[37] LTS estimates were obtained using Eric Blankmeyer's LTS routine in RATS (version 4.3). SHAZAM (version 7.0) or TSP (version 4.4) were used for all other computations.

[38] Their results for life expectancy are, however, much less precise, and hence less compelling, than for infant and child mortality.

it is necessary to obtain instruments that are correlated with the regressors, but uncorrelated with the country-specific error term. Bils and Klenow (1998) suggest that climate variables can be used as suitable instruments for education and health. Details of these climate variables are given in the notes to Table 4 and in Appendix 1. Hall and Jones (1999) argue that measures of Western European influence will be uncorrelated with the country-specific error term, but correlated with social infrastructure which, in turn, is correlated with capital accumulation and educational attainment. We use two of Hall and Jones' measures: distance from the equator[39] and the proportion of the population speaking a Western European language as a first language. We also include regional dummies as instruments (as do Hall and Jones) and Catholicism and Muslim dummies, which allow for the effects of religion on institutional quality, including government performance, (La Porta *et al.*, 1999) and the human capital variables.

The results obtained using different combinations of the instruments are reported in Table 4. To check on the correlation between the residuals and the instruments we calculated a general misspecification test for instrumental variables estimation of over-identified models, denoted *SARG*. We also calculated a Hausman test (denoted *HAUS*) of the consistency of the OLS estimates by comparison with 2SLS based on a presumed-valid instrument set. Due to a lack of data for the instruments for some countries the sample size is reduced to 60. Column (i) of Table 3 reports the OLS results obtained for eq. (14) for the reduced sample. For this sample α is not statistically significant.[40] The Sargan test results suggest that the instrument sets for columns (ii) and (iii) are valid, but not that for the restricted model in column (iv). The Hausman test results suggest that the OLS estimates are not significantly affected by endogeneity. In the 2SLS results, the elasticity with respect to female education is consistently positive and significant. The elasticity with respect to male schooling is again negative but, unlike the OLS results, is now statistically significant. The restriction $\beta_f = -\beta_m$ is not rejected for the results in columns (ii) to (iv) (with asymptotic t-values of -0.990, -1.273 and -0.555 for columns (ii), (iii) and (iv) respectively). These results imply a significant long-run effect of the educational gender gap on output per worker, but no additional effects from either of the separate levels of schooling. Because of the difficulties in obtaining suitable instruments in cross-section studies it is easy to overstate evidence concerning direction of causation. However, our finding that, other things equal, higher levels of female education lead to higher levels of labour

[39] Hall and Jones argue that as Western Europeans tended to settle in temperate regions, there will be a positive relationship between distance from the equator and Western European influence.

[40] This appears to reflect the sensitivity of the estimates of α to the sample of countries selected. However, it may also be the case that, for all the sets of estimates, some of the effects of population growth and physical capital investment are picked up by the coefficient on female education, if increases in female education reduce fertility and raise steady-state levels of output and, hence, physical capital stock per worker (Temple, 2000).

Table 4 OLS and 2SLS estimates of the steady-state equation for GDP per worker
Dependent variable: average log of income per worker 1960–1990

Variable	(i) OLS	(ii) 2SLS	(iii) 2SLS	(iv) 2SLS
Unrestricted model				
$\ln(s_k)$	0.005	0.274	0.281	0.080
	(0.04)	(0.86)	(0.84)	(0.30)
$\ln(n+g+\delta)$	−0.127	−1.330	−1.449†	−1.331†
	(−0.43)	(−1.62)	(−1.71)	(−1.85)
$\ln(ef)^*$	0.497**	0.830**	0.822**	0.759**
	(3.88)	(3.52)	(3.37)	(3.62)
$\ln(em)^*$	−0.273	−0.943*	−1.014*	−0.623
	(−1.54)	(−1.98)	(−2.07)	(−1.60)
$\ln(x)^*$	0.418*	0.075	0.044	0.011
	(2.29)	(0.15)	(0.09)	(0.02)
$\ln(A)$	0.810**	0.805**	0.893**	0.858**
	(12.52)	(3.32)	(3.89)	(5.19)
R^2 adjusted	0.918	0.868	0.857	0.889
SARG		6.249 (4)	7.547 (6)	14.910†(8)
HAUS		5.384 (6)	6.337 (6)	6.035 (6)
Restricted model				
$\ln(s_k) - \ln(n+g+\delta)$	0.016	0.337	0.352	0.170
	(0.14)	(1.09)	(1.08)	(0.69)
$\ln(ef)^*$	0.484**	0.710**	0.690**	0.621**
	(3.96)	(3.33)	(3.09)	(3.41)
$\ln(em)^*$	−0.271	−1.069*	−1.159*	−0.773*
	(−1.54)	(−2.34)	(−2.44)	(−2.17)
$\ln(x)^*$	0.447**	0.564†	0.590†	0.603*
	(2.61)	(1.70)	(1.69)	(2.08)
$\ln(A)$	0.805**	0.758**	0.838**	0.769**
	(11.91)	(3.22)	(3.74)	(5.21)
R^2 adjusted	0.919	0.873	0.859	0.902
F	0.174	1.711	1.971	2.978
SARG		8.352 (5)	9.735 (7)	20.361*(9)
HAUS		3.866 (5)	4.389 (5)	3.123 (5)
N	60	60	60	60
Implied elasticities				
Implied α	0.016	0.252	0.260	0.145
	(0.14)	(1.45)	(1.46)	(0.81)
Implied β_f	0.477**	0.531**	0.510**	0.530**
	(3.65)	(2.78)	(2.61)	(2.80)
Implied β_m	−0.267	−0.800**	−0.857**	−0.661*
	(−1.57)	(−2.71)	(−2.82)	(−2.39)
Implied ψ	0.440*	0.422*	0.436	0.515†
	(2.21)	(2.14)	(1.44)	(1.72)

Notes: Asymptotic t-statistics are given in parentheses. *SARG* is Sargan's misspecification test for 2SLS estimation and *HAUS* is a test of the consistency of the OLS estimates. Degrees of freedom for asymptotically chi-squared distributed *SARG* and *HAUS* tests are given in parentheses. For other notes see Table 1.

Instrument sets: Column (ii): the ratio of rain in the heaviest month to the average, the log of the average high temperature in April (October for the Southern hemisphere), the difference between the highest monthly high temperature and the lowest monthly low, the ratio of country border made up of coast, the distance from the equator (data from Bils and Klenow, 1998), the influence of Western European languages (data from Hall and Jones, 1999), and East Asia, Latin America, Sub-Saharan Africa, and OECD dummies; column (iii): as for column (ii) plus average rainfall and the altitude of the most populous city; column (iv) as for column (iii) plus dummy variables (=1) for countries in which a majority of the population is Catholic or Muslim.

productivity, and the gender gap has an adverse effect, is robust for the instrument sets examined.[41]

8. Conclusion

This paper set out to examine the implications of educational gender gaps on development and, in particular, whether, in the long run, increasing female schooling leads to higher levels of labour productivity across countries. Although a substantial literature using micro data supports this view, the existing cross-country growth literature, using macro data, is inconclusive. Much of this cross-country literature is somewhat *ad hoc*, making the coefficients on female education difficult to interpret.

In this paper, we derive and estimate a neoclassical growth model that includes female and male education as separate explanatory variables. The model can be reparameterised so that education enters as a gender gap. The interpretation of the coefficient on the gender gap variable depends crucially on what other education variables (male or female) are included in the equation.

Drawing on recent developments in regression limit theory for non-stationary panel data, we estimate the cross-section average long-run effect of female and male education on output per worker using long time averages of the data. Further work on the implications for the asymptotic theory of cross-section dependence and evidence on finite-sample properties when using long period averages to estimate long-run effects is clearly desirable. However, it does provide a viable alternative to the conventional approach of extracting key elasticities from transitional growth equations, which has some serious problems. Our empirical results suggest that female education has a statistically significant positive effect on labour productivity. The role of male education is less clear; the relevant point estimate and implied elasticity are usually negative but their statistical significance is sensitive to the method of estimation and the choice of other variables included.

Our results on the role of female education are robust to a variety of sensitivity analyses including testing for the effect of influential observations and using instrumental variables to allow for the possibility of simultaneity bias. The difference in our results compared to, for example, Barro and Lee (1994) appears to be due to their use of base-period values of human capital stock measures, which is not consistent with our explicit model. Our coefficient estimates for female schooling are consistent with the estimated effects of education obtained in the labour economics literature, using micro data. For the OLS and, particularly, the 2SLS

[41] The marginal significance level of the elasticity with respect to life expectancy varies across instrument sets, and is significant at the 5% level in column (ii) only. If life expectancy is deleted from the models in columns (ii) to (iv), the results for β_f and β_m (including non-rejection of $\beta_f = -\beta_m$) are qualitatively unchanged, although for the instrument set in (ii) and (iii) α is significant at the 5% level with a point estimate approximately equal to 0.32. For column (iv), α is not statistically significant but, again, this set of instruments is rejected by the Sargan tests.

estimates, the results are consistent with a significant long-run effect of the educational gender gap on output per worker, but no additional effects from either of the separate levels of schooling.

The title of this paper asks 'are educational gender gaps a brake on economic development?' Any conclusion must be tempered by acknowledging that disentangling patterns of significance and causation from highly collinear, interrelated cross-country data of varying quality is fraught with difficulties. However, our results suggest that educational gender gaps are an impediment to economic development. The World Bank's claims about the importance of female education do not seem misplaced.

Acknowledgements

Previous versions of this paper have been presented at the Econometric Society Australasian Meeting (Canberra, July 1998), the European Economic Association Congress (Berlin, September 1998), the New Zealand Association of Economists Conference (Wellington, September 1998), the Economic Society of Australia Conference (Sydney, September 1998), a New Zealand Econometric Study Group Meeting (Hamilton, February 1999), the Royal Economic Society Conference (St Andrews, July 2000), and in seminars at the Universities of Dundee, Nottingham, and Otago. We are grateful to two referees, Christopher Adam (Associate Editor), Steve Dowrick, Stephan Klasen, Peter Phillips and conference participants for helpful comments and suggestions, to Mark Bils for providing the climate data used in the instrumental variables estimation, and to Eric Blankmeyer for advice on his LTS procedure. We are also grateful for financial support from the Marsden Fund, administered by the Royal Society of New Zealand, and from a University of Otago Research Grant. 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

Anand, S. and Ravallion, M. (1993). 'Human development in poor countries: on the role of private incomes and public services', *Journal of Economic Perspectives*, **7,** 133–50.

Baltagi, B.H. and Griffin, J.M. (1984). 'Short and long run effects in pooled models', *International Economic Review*, **25,** 631–45.

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

Barro, R.J. (1996a). 'Inflation and growth', *Federal Reserve Bank of St. Louis Review*, **78,** 153–69.

Barro, R.J. (1996b). 'Democracy and growth', *Journal of Economic Growth*, **1,** 1–27.

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

Barro, R.J. (1998). 'Human capital and growth in cross-country regressions', manuscript, Harvard University, (http://hassler-j.iies.su.se/conferences/papers/barro.pdf).

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.

Behrman, J.R. and Deolalikar, A.B. (1988). 'Health and nutrition', in H. Chenery and T.N. Srinivasan (eds), *Handbook of Development Economics, Volume I*, North-Holland, Amsterdam.

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

Benavot, A. (1989). 'Education, gender, and economic development: a cross-national study', *Sociology of Education*, **62,** 14–32.

Bils, M. and Klenow, P.J. (1998). 'Does schooling cause growth or the other way around?', Working Paper 6393, NBER.

Birdsall, N., Ross, D., and Sabot, R. (1997). 'Education, growth and inequality', in N. Birdsall and F. Jaspersen (eds), *Pathways to Growth: Comparing East Asia and Latin America*, Johns Hopkins University Press, Washington, DC.

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.

De la Fuente, A. and Doménech, R. (2000). 'Human capital in growth regressions: how much difference does data quality make?', Discussion Paper No. 2466, Centre for Economic Policy Research, London.

Dollar, D. and Gatti, R. (1999). 'Gender inequality, income, and growth: are good times good for women?', Policy Research Report on Gender and Development Working Paper Series, No. 1, The World Bank, Washington, DC.

Forbes, K.J. (2000). 'A reassessment of the relationship between inequality and growth', *American Economic Review*, **90,** 869–87.

Gemmell, N. (1996). 'Evaluating the impacts of human capital stocks and accumulation on economic growth: some new evidence', *Oxford Bulletin of Economics and Statistics*, **58,** 9–28.

Hall, R.E. and Jones, C.I. (1999). 'Why do some countries produce so much more output per worker than others?', *Quarterly Journal of Economics*, **114,** 83–115.

Hendry, D.F. (1995). *Dynamic Econometrics*, Oxford University Press, Oxford.

Hill, M.A. and King, E.M. (1993). 'Women's education in developing countries: an overview', in E.M. King and M.A. Hill (eds),*Women's Education in Developing Countries: Barriers, Benefits, and Policies,* Johns Hopkins University Press, Baltimore, MD.

Hill, M.A. and King, E.M. (1995). 'Women's education and economic well-being', *Feminist Economics*, **1,** 21–46.

Islam, N. (1995). 'Growth empirics: a panel data approach', *Quarterly Journal of Economics*, **110,** 1127–70.

Jones, C. (1997). 'Convergence revisited', *Journal of Economic Growth*, **2,** 131–53.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H., and Lee, T.-C. (1988). *Introduction to the Theory and Practice of Econometrics*, John Wiley, New York, second edition.

Klasen, S. (1999). 'Does gender inequality reduce growth and development? Evidence from cross-country regressions', Policy Research Report on Gender and Development Working Paper Series, No. 2, The World Bank, Washington, DC.

Klenow, P. and Rodriguez-Clare, A. (1997). 'The neoclassical revival in growth economics: has it gone too far?', in B. Bernanke and J. Rotemberg (eds), *1997 NBER Macroeconomics Annual*, MIT Press, Cambridge, MA.

Knight, M., Loayza, N., and Villanueva, D. (1993). 'Testing the neoclassical theory of economic growth', *IMF Staff Papers*, **40,** 512–41.

Knowles, S. and Owen, P.D. (1995). 'Health capital and cross-country variation in income per capita in the Mankiw-Romer-Weil model', *Economics Letters*, **48,** 99–106.

Knowles, S. and Owen, P.D. (1997). 'Education and health in an effective-labour empirical growth model', *Economic Record*, **73,** 314–28.

Krueger, A.B. and Lindahl, M. (1998). 'Education for growth: why and for whom', manuscript, Department of Economics, Princeton University, Princeton, NJ.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. (1999). 'The quality of government', *Journal of Law, Economics, & Organization*, **15,** 222–82.

Lee, K., Pesaran, M.H., and Smith, R. (1997). 'Growth and convergence in a multi-country empirical stochastic Solow model', *Journal of Applied Econometrics*, **12,** 357–92.

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.

Mankiw, N.G., Romer, D., and Weil, D.N. (1992). 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics*, **107,** 407–37.

Mincer, J. (1974). *Schooling, Earnings and Experience*, Columbia University Press, New York.

Moon, H.R. and Phillips, P.C.B. (1998). 'A reinterpretation of the Feldstein-Horioka regressions from a nonstationary panel viewpoint', mimeo, Yale University, New Haven, CT.

Perotti, R. (1996). 'Growth, income distribution, and democracy: what the data say', *Journal of Economic Growth*, **1,** 149–87.

Pesaran, M.H. and Smith, R. (1995). 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of Econometrics*, **68,** 79–113.

Phillips, P.C.B. and Moon, H.R. (1999). 'Linear regression limit theory for nonstationary panel data', *Econometrica*, **67,** 1057–111.

Phillips, P.C.B. and Moon, H.R. (2000). 'Nonstationary panel data analysis: an overview of some recent developments', *Econometric Reviews*, **19,** 263–86.

Pritchett, L. (1996). 'Where has all the education gone?', Working Paper No.1581, Policy Research Department, The World Bank, Washington, DC.

Pritchett, L. and Summers, L.H. (1996). 'Wealthier is healthier', *Journal of Human Resources*, **31,** 841–68.

Psacharopoulos, G. (1994). 'Returns to investment in education: a global update', *World Development*, **22,** 1325–43.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*, John Wiley, New York.

Schultz, T.P. (1988). 'Education investments and returns', in H. Chenery and T.N. Srinivasan (eds), *Handbook of Development Economics*, *Volume I*, North-Holland, Amsterdam.

Schultz, T.P. (1995). 'Investments in schooling and health of women and men: quantities and returns', in T.P. Schultz (ed.), *Investment in Women's Human Capital*, University of Chicago Press, Chicago, IL.

Sen, A. (1998). 'Mortality as an indicator of economic success and failure', *Economic Journal*, **108,** 1–25.

Sims, C.A. (1996). 'Commentary [on R. J. Barro, Inflation and growth]', *Federal Reserve Bank of St. Louis Review*, **78,** 173–8.

Solow, R.M. (1956). 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70,** 65–94.

Stokey, N. (1994). 'Comments on Barro and Lee', *Carnegie-Rochester Conference Series on Public Policy*, **40,** 47–57.

Subbarao, K. and Raney, L. (1995). 'Social gains from female education: a cross-national study', *Economic Development and Cultural Change*, **44,** 105–28.

Swan, T.W. (1956). 'Economic growth and capital accumulation', *Economic Record*, **32,** 334–61.

Temple, J. (1998a). 'Equipment investment and the Solow model', *Oxford Economic Papers*, **50,** 39–62.

Temple, J. (1998b). 'Robustness tests of the augmented Solow model', *Journal of Applied Econometrics*, **13,** 361–75.

Temple, J. (1999a). 'The new growth evidence', *Journal of Economic Literature*, **37,** 112–56.

Temple, J. (1999b). 'A positive effect of human capital on growth', *Economics Letters*, **65,** 131–4.

Temple, J. (2000). 'Growth effects of education and social capital in the OECD countries', Economics Department Working Papers No. 263, OECD.

Topel, R. (1999). 'Labor markets and economic growth' in O. Ashenfelter and D. Card (eds) *Handbook of Labor Economics, Volume 3C*, North-Holland, Amsterdam.

Wheeler, D. (1980). 'Basic needs fulfillment and economic growth: a simultaneous model', *Journal of Development Economics*, **7,** 435–51.

Wolfensohn, J. (1995). 'Women and the transformation of the 21st century', Address to the Fourth UN Conference on Women, Beijing.

World Bank (1993). *World Development Report* 1993: *Investing in Health*, Oxford University Press, New York.

World Bank (1994) *Enhancing Women's Participation in Economic Development*, The World Bank, Washington DC.

Appendix 1: Data Sources

Y/L Real GDP per worker, from the Penn World Tables version 5.6 (PWT 5.6) (http://www.nber.org/pwt56.html). Annual data were logged then averaged over the period 1960 to 1990.

s_k Ratio of real physical capital investment to real GDP from PWT 5.6. Annual data were logged then averaged over the period 1960 to 1990.

$(n + g + \delta)$ Adjusted growth rate of the labour force. The number of workers for each year (1960–1990) was calculated using income *per capita*, income per worker and population data, from PWT 5.6. The growth rate of the workforce (n) was calculated for each year.

(g+d) was assumed to equal 5%. (n+g+d) was logged then averaged over the period 1960 to 1990.

ef Average number of years of schooling attained by the female population aged 15 years and over. The data, taken from Barro and Lee (1996) and available at five-year intervals, were logged then averaged over the period 1960 to 1990.

em Average number of years of schooling attained by the male population aged 15 years and over. The data, taken from Barro and Lee (1996) and available at five-year intervals, were logged and then averaged over the period 1960 to 1990.

x The shortfall in life expectancy at birth (*LE*) from 85 years. The data used are annual (1965 to 1990) from the World Bank's *Social Indicators of Development* 1991–92 diskette. The shortfall is calculated as $\ln(x) = -\ln(85 - LE)$. The data were logged then averaged over the period 1965 to 1990.

The data for the additional Barro control variables (row (b) of Table 3) and the Barro and Lee (1993) education data (row (a) of Table 3) are from the Barro-Lee (1994) data set (http://www.nber.org).

The data for Hall and Jones' measure of technical efficiency are from Hall and Jones (1999) (http://www.stanford.edu/~chadj/HallJones400.asc).

The climate data (the ratio of rain in the heaviest month to the average, the log of the average high temperature in April (October for the Southern hemisphere), the difference between the highest monthly high temperature and the lowest monthly low, the ratio of country border made up of coast, average rainfall, altitude of the most populous city) are from Bils and Klenow (1998).

The data on the influence of Western European languages and distance from the equator are from Hall and Jones (1999) (http://www.stanford.edu/~chadj/HallJones400.asc).

Dummy variables for religion (Catholic and Muslim) take a value of one if the majority of the population is of the relevant religion and zero otherwise. The data on religion are from the travel finder internet site (http://www.travelfinder.com/wfacts2/country.html).

Appendix 2

Table A2 Educational gender gaps across countries

Country	Average total years of female schooling	Average total years of male schooling	Gender gap expressed as a ratio
Africa			
Algeria	1.49	2.87	0.52
Cameroon	1.57	2.97	0.53
Central African Republic	0.56	1.67	0.34
Ghana	1.60	3.95	0.41
Kenya	1.75	3.42	0.51
Malawi	1.56	3.22	0.49
Mali	0.34	0.88	0.39
Mauritius	3.76	5.21	0.72
Mozambique	0.29	1.10	0.26
Senegal	1.39	2.58	0.54
South Africa	4.38	4.59	0.95
Togo	0.78	2.30	0.34
Tunisia	1.50	2.93	0.51
Uganda	0.85	2.01	0.42
Zambia	2.44	4.79	0.51
Zimbabwe	1.73	2.79	0.62
North and Central America			
Canada	9.97	9.99	1.00
Costa Rica	4.76	4.78	1.00
Dominican Republic	3.38	3.60	0.94
El Salvador	2.56	3.13	0.82
Guatemala	1.91	2.44	0.79
Honduras	2.50	3.18	0.79
Jamaica	3.90	3.42	1.14
Mexico	3.87	4.70	0.83
Nicaragua	2.86	3.15	0.91
Panama	5.77	5.67	1.02
Trinidad & Tobago	5.89	6.01	0.98
United States of America	10.27	10.30	1.00
South America			
Argentina	6.46	6.53	0.99
Bolivia	3.99	5.78	0.69
Brazil	3.22	3.20	1.01
Chile	5.78	6.00	0.96
Colombia	3.91	3.88	1.01
Ecuador	4.46	4.84	0.92
Paraguay	4.16	4.67	0.89
Peru	4.30	5.46	0.79
Uruguay	6.20	5.93	1.04
Venezuela	4.04	4.21	0.96

(*continued*)

Table A2—(*continued*)

Country	Average total years of female schooling	Average total years of male schooling	Gender gap expressed as a ratio
Asia			
Bangladesh	0.68	2.04	0.33
Hong Kong	5.90	8.24	0.72
India	1.62	3.92	0.41
Indonesia	2.45	3.71	0.66
Israel	8.11	8.87	0.91
Japan	7.81	8.44	0.93
Jordan	2.89	4.93	0.59
Korea	5.74	7.91	0.73
Malaysia	3.29	5.64	0.58
Pakistan	0.98	2.73	0.36
Philippines	5.53	5.78	0.96
Singapore	4.43	6.13	0.72
Sri Lanka	4.47	5.44	0.82
Syria	1.92	4.12	0.47
Taiwan	4.95	7.35	0.67
Thailand	4.01	4.99	0.80
Europe			
Austria	5.20	6.60	0.79
Belgium	8.12	8.64	0.94
Denmark	9.93	10.48	0.95
Finland	8.73	8.67	1.01
France	5.40	5.73	0.94
Greece	5.27	7.26	0.73
Ireland	7.25	7.06	1.03
Italy	5.24	6.06	0.87
Netherlands	7.23	7.65	0.94
Norway	6.52	7.03	0.93
Portugal	2.71	3.50	0.78
Spain	5.07	5.52	0.92
Sweden	8.59	9.03	0.95
Switzerland	7.56	8.32	0.91
Turkey	2.01	3.62	0.56
United Kingdom	8.09	8.03	1.01
Australasia			
Australia	9.67	10.28	0.94
New Zealand	10.62	10.96	0.97
Papua New Guinea	1.17	1.91	0.61

Notes: The data are from Barro and Lee (1996) and are averages for the period 1960 to 1990 for the population aged 15 years and over. The gender gap = the average total years of female schooling divided by the average total years of male schooling.