JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **21**: 111–132 (2006)
Published online 9 June 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jae.813

ECONOMIC DEVELOPMENT AND THE RETURN TO HUMAN CAPITAL: A SMOOTH COEFFICIENT SEMIPARAMETRIC APPROACH

THEOFANIS P. MAMUNEAS,[a] ANDREAS SAVVIDES[b] AND THANASIS STENGOS[c]*

[a] *Department of Economics, University of Cyprus, Cyprus*
[b] *Department of Economics, Oklahoma State University, USA*
[c] *Department of Economics, University of Guelph, Canada*

SUMMARY

This paper investigates the impact of human capital on the process of economic growth by allowing the contribution of traditional inputs (capital and labour) as well as that of human capital to vary both across countries and time. The former is accomplished by constructing an index of TFP growth for traditional inputs, while the latter through semiparametric methods. We derive estimates of the output elasticity and social return to human capital for 51 countries at various stages of economic development.

1. INTRODUCTION

The contribution of human capital to the growth of income is controversial. At the micro level, there is consistent evidence that education raises incomes significantly (commonly referred to as Mincerian wage regressions). Evidence at the macro level has been mixed. Studies such as Barro (1991), Bils and Klenow (2000), Mankiw *et al.* (1992) and others use enrollment rates and find a positive and significant contribution for human capital to the growth of output (GDP). Benhabib and Spiegel (1994), Kyriacou (1991), Lau *et al.* (1991) and Pritchett (2001), however, find an insignificant or even negative contribution for the stock of human capital (mean years of schooling). The estimated effect for human capital does not hinge on the way it is defined (as stock or flow).[1] For example, Barro and Sala-i-Martin (1995) find the impact of enrollment rates to be insignificant while mean years of schooling has a positive and significant effect on economic growth. Recent work has focused on the possibility of multiple regimes and nonlinearities in the human capital–growth process. Durlauf and Johnson (1995) and Masanjala and Papageorgiou (2002) use the regression-tree and the threshold-regression methodology to show the existence of multiple regimes. Kalaitzidakis *et al.* (2001) use semiparametric techniques and find that there are substantial nonlinearities in the growth–human capital relationship that linear models are unable to detect. Kourtellos (2003) also uses a semiparametric coefficient model to study a local generalization of the Solow model in the spirit of Durlauf *et al.* (2001). He examines two possibilities: first, the parameters of the Solow growth model are a smooth unknown function of

* Correspondence to: Thanasis Stengos, Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. E-mail: tstengos@uoguelph.ca

[1] Early studies tended to use a flow measure (enrollment rates) while more recent studies have used a stock measure (mean years of schooling derived from cumulating past enrollment rates).

Received 22 July 2003
Revised 7 June 2004

the literacy rate and second, a function of life expectancy (interpreted as a measure of the quality of human capital).

The main purpose of this paper is to estimate a general model of the economic growth process by allowing the contribution of both 'traditional' inputs (capital and labour) and human capital to vary across both countries and time. In the first place we calculate an index of total factor productivity (TFP) growth that contains only the contribution of the 'traditional' inputs (capital and labour). This index uses measured data on the share of labour and capital in total output and thus eschews the estimation of a production function with the implicit assumption that the output contribution of capital and labour is constant. This is an important attribute because all previous studies (including ours) contain countries at widely differing stages of economic development, obviating the need to relax the assumption of a constant contribution both intra- and inter-temporally. Next, we use this index to evaluate the impact of human capital growth on the growth of TFP. We accomplish this via a semiparametric smooth coefficient model that allows for the effect of human capital on economic growth to be nonlinear. Theoretical work by Azariadis and Drazen (1990) has demonstrated that the way in which human capital impacts the growth of output differs across countries. In particular, threshold externalities may exist as a result of attaining a 'critical mass' in human capital and, therefore, economies that are similar in terms of technology or preferences may display substantial differences in growth rates if they lie on either side of the threshold. In Durlauf (1993), local technological spillovers generate multiple growth equilibria characterized by differences in the way in which the traditional inputs are employed in aggregate production.

An important contribution of our work is the construction of estimates of the elasticity of output with respect to human capital that display a rather large variation across countries, casting doubt on the equality assumption in the literature. Moreover, for a number of the developing economies in our sample, human capital (mean years of schooling) has no significant effect on aggregate output, a possibility mentioned in the literature (e.g. Pritchett, 2001) but not explored in a systematic fashion empirically.

In the next section we present our methodology. We construct a traditional index of TFP growth that does not rely on the estimation of a production function with the concomitant assumption of constant capital and labour contributions across countries and time. In the third section we present the smooth coefficient semiparametric model that allows the effect of human capital on TFP growth to be nonlinear. This methodology enables us to derive country (and time)-specific estimates of output elasticities and social rates of return for human capital. The final section concludes the paper.

2. METHODOLOGY

The treatment of human capital in cross-country (macro) studies falls into three main categories: (a) a production function is combined with equations for the accumulation of reproducible factors to arrive at equations for the steady state of per capita income and the transition process to the steady state (e.g. Mankiw *et al.*, 1992 and the substantial literature spawned by this study); (b) a production function that includes human capital as either an input or as a determinant of total factor productivity is estimated directly (e.g. Benhabib and Spiegel, 1994; Edwards, 1998; Pritchett, 2001); and (c) specific values are assigned to the coefficients of the production function (the shares of inputs) to arrive at an estimate of the residual which is then evaluated, along with the factors of production (including human capital), for their contribution to differences in the level or growth of per capita GDP (e.g. Hall and Jones, 1999; Klenow and Rodríguez-Clare, 1997). Most of these studies are based on very restrictive assumptions: a Cobb–Douglas production function with Hicks-neutral technology. By contrast, our study relies on a general framework. We assume that a general production function describes the technology of country i at time t as follows:

$$Y = f(K, E, H, t) \tag{1}$$

where Y, K and E represent the amounts of total output, physical capital and effective or human-capital augmented labour, respectively, H is average human capital (mean years of schooling per person of working age) and t is a technology index measured by the time trend.

Total differentiation of (1) with respect to time and division by Y yields

$$\widehat{Y} = \widehat{A} + \varepsilon_K \widehat{K} + \varepsilon_E \widehat{E} + \varepsilon_H \widehat{H} \tag{2}$$

where $(\widehat{\ })$ denotes a growth rate, $\widehat{A} = (\partial f/\partial t)/Y$ is the exogenous rate of technological change and $\varepsilon_Q = \partial \ln Y/\partial \ln Q$ $(Q = K, E, H)$ denotes output elasticity. The last term in (2) measures the externality effect on human capital accumulation. This effect is emphasized by recent endogenous growth models initiated by Lucas (1988). Assuming a perfectly competitive theory of distribution, the output elasticities of effective labour and capital should be equal to the observed income shares of labour, s_{YL} and capital, s_{YK}.[2] Equation (2), however, is not useful for empirical purposes because the growth rate of effective labour $\widehat{E}$ is not observable.

Assuming that the effective labour input is a function of the labour force and average human capital, or

$$E = g(L, H) \tag{3}$$

we can decompose $\widehat{E}$ as

$$\widehat{E} = \eta_L \widehat{L} + \eta_H \widehat{H} \tag{4}$$

where η_L and η_H are effective labour elasticities with respect to labour and average human capital, respectively. Substituting (4) in (2) we have:

$$\widehat{Y} = \widehat{A} + \varepsilon_K \widehat{K} + \varepsilon_E \eta_L \widehat{L} + (\varepsilon_E \eta_H + \varepsilon_H)\widehat{H} \tag{5}$$

By contrast with (2), the last term in parentheses in (5) measures the total effect of human capital, while the output elasticity of raw labour is $\varepsilon_E \eta_L$. Direct estimation of (5) corresponds to the well-known growth accounting methodology.

As mentioned previously and by several authors (e.g. Temple, 1999), an important criticism of estimating a model such as (5) is the assumption that the contribution of inputs is the same across countries and time so the estimated parameters represent an 'average' contribution. In the remainder of the paper we address this issue through an alternative specification that accounts for differing contributions of all productive inputs. In the first instance, we construct an index of TFP growth for our panel that contains only the traditional inputs. This index allows the contribution

[2] The income share of labour can be defined in two ways: the income share of effective labour (E) or the income share of 'traditional' or workforce labour (L). These two are equal since the value of labour (the numerator of the income share) is the same independently of the definition of the labour input as L or E (the corresponding price and quantity indices, however, will differ).

of capital and labour to differ both intra- and inter-temporally and to be dictated by the data. We define a Törnqvist index of TFP growth for country i $(1, \ldots, N)$ in year t $(1, \ldots, T)$ as follows:

$$\widehat{TFP}_{it} = \widehat{Y}_{it} - s_{Eit}\widehat{L}_{it} - s_{Kit}\widehat{K}_{it} \tag{6}$$

where $s_{E_{it}} = 0.5(S_{L_{it}} + S_{L_{it-1}})$ and $s_{K_{it}} = 0.5(S_{K_{it}} + S_{K_{it-1}})$ are weighted averages of the cost shares of labour and physical capital ($S_{L_{it}}$ and $S_{K_{it}}$, defined in Appendix A) and $\widehat{Q}_{it} = \ln Q_{it} - \ln Q_{it-1}(Q = Y, L, K)$.[3] This measure of TFP contains the component of output growth that cannot be explained by the growth of 'traditional' inputs (K and L). This index will be an exact index of exogenous technological change under certain conditions (see Diewert, 1976) and as long as the production function contains only the traditional inputs. The production function underlying this index is a general translog. It should be stressed, however, that if human capital enters the production function as in (1), this index is a biased index of technological change and, therefore, variations in $\widehat{H}$ will affect TFP growth as defined in (6). In what follows we present a methodology for estimating this effect.

Subscripting equation (5) by country and year (it), taking a discrete approximation of the continuous growth rates and adding (6) to it we have:

$$\begin{aligned}\widehat{TFP}_{it} = \widehat{A}_{it} &+ [(\varepsilon_{Kit} - s_{Kit})\widehat{K}_{it} + (\varepsilon_{Eit}\eta_L - s_{Eit})\widehat{L}_{it}] \\ &+ (\varepsilon_{Eit}\eta_{Hit} + \varepsilon_{Hit})\widehat{H}_{it}\end{aligned} \tag{7}$$

where the first term ($\widehat{A}_{it}$) is the exogenous rate of technological change, and the final term in parentheses is the total contribution of human capital. The latter is made up of two components: the first is the direct or private effect of human capital and the second is the indirect or externality effect. The term in brackets is the scale effect. Under constant returns to scale, output elasticities will be equal to the cost shares and the term in brackets will be equal to zero if $\eta_L = 1$.[4] It can be shown that the first-order conditions of standard cost minimization with respect to physical capital and labour (taking average human capital as given) imply that

$$\varepsilon_{jit} = \rho s_{jit}, \quad j = K, E \tag{8}$$

where $\rho = \varepsilon_{CY}^{-1}$ is the elasticity of returns to scale of capital and labour and $\varepsilon_{CY} = (\partial C/\partial Y)/(Y/C)$ is cost flexibility.

Of central importance to our study is the final term in (7) that captures the contribution of human capital to aggregate production. We model this as a general unknown function $\theta(\cdot)\widehat{H}_{it}$. Details on the specification and estimation of the $\theta(\cdot)$ function will be provided in the following section. Using the above formulation for $\theta(\cdot)$ and (8), (7) can be written as

$$\widehat{TFP}_{it} = \widehat{A}_{it} + \alpha\widehat{M}_{it} + \theta(\cdot)\widehat{H}_{it} \tag{9}$$

[3] The cost shares of effective and 'traditional' labour are the same independently of how we define labour because labour cost should be the same. Thus $s_{Eit} \equiv s_{L_{it}}$, where E and L denote, as before, an index of effective and 'traditional' labour input.

[4] The condition $\eta_L = 1$ will be true if (3) is linear homogeneous in L or the production function for effective labour can be written as $g(L, H) = L\varphi(H)$, where $\varphi(H)$ is any function of H. Many popular models belong to this class of effective labour production function. For example, Lucas (1988) postulates that $g(L, H) = LH$. Another possibility is a Mincerian-type production function $g(L, H) = Le^{\phi(H)}$ assumed by Hall and Jones (1999).

where $\alpha = (\rho - 1)$ and $\widehat{M}_{it} = s_{K_{it}}\widehat{K}_{it} + s_{Eit}\widehat{L}_{it}$. Estimation of (9) allows testing the hypothesis of nonconstant returns to scale in capital and labour ($\alpha \neq 0$). Moreover, it allows human capital to influence TFP growth in a nonlinear fashion.

In the following section the nonlinear aspects of human capital growth are investigated systematically via semiparametric estimation. Equation (9) forms the basis of our empirical analysis. It is important to note that the form and interpretation of (9) depends on the definition of the TFP growth index in (6). Our objective is to identify the total effect of human capital (direct plus indirect) under minimal assumptions. Therefore in (6) we define an index that can be constructed from observable data and can yield estimates of the total effect. For instance, instead of (6) we can define a TFP growth index by subtracting from output growth the (cost-share) weighted growth rates of effective labour and physical capital, $\widehat{TFP}_{it} = \widehat{Y}_{it} - s_{Eit}\widehat{E}_{it} - s_{K_{it}}\widehat{K}_{it}$, or by subtracting from output growth the (cost-share) weighted growth rates of labour and capital where the labour weight is based on the cost share of unskilled or raw labour $s^b_{L_{it}}$, $\widehat{TFP}_{it} = \widehat{Y}_{it} - s^b_{L_{it}}\widehat{L}_{it} - s_{K_{it}}\widehat{K}_{it}$. These alternative definitions of the TFP index will not meet our objectives. Using these alternative indices would require information on effective labour or the raw labour cost share, neither of which is directly observable. It could be argued, however, that these can be obtained by making assumptions about the private contribution of human capital or by using estimates from micro studies. Clearly, if these assumptions or estimates are not correct they may bias the total effect of human capital. In summary, defining a TFP growth index as in (6) that is based on observable data appears to us appropriate.

3. ESTIMATION AND EMPIRICAL RESULTS

3.1. Estimation Method

In order to estimate equation (9) we model $\widehat{A}_{it}$ (exogenous growth of technological change) as a function of country- and year-specific dummy variables. Country-specific dummies (D_i) capture idiosyncratic exogenous technological change and time-specific dummies (D_t) capture the procyclical behaviour of TFP growth. In alternative specifications we include political and economic freedom and a country's trade orientation as additional determinants of TFP growth. The rationale for this is that increased exposure to international trade promotes technology absorption and boosts productivity. Edwards (1998) provides evidence linking outward orientation (Z_1) with TFP growth. The role of institution building as a determinant of long-run economic performance has received considerable attention recently. Rodrik (2000) discusses five types of institutions: property rights, regulatory institutions, institutions for macroeconomic stabilization, institutions for social insurance and institutions of conflict management. He argues that building institutions can be thought of as a form of technology transfer that allows increased productivity. Participatory democracy is a meta-institution that helps build better institutions. He provides evidence that participatory democracy improves economic performance both in terms of higher long-run growth rates and short-term stability. We include two measures of participatory democracy: an index of political freedoms (Z_2) and civil freedoms (Z_3).[5]

As for the unknown function $\theta(\cdot)$, we estimate two alternative specifications: (i) it depends only on the level of human capital and (ii) in addition to the level of human capital it also depends

[5] Details on the measurement of the political and civil freedom indexes and outward orientation are provided in Appendix A.

on other economy characteristics Ω. In our case Ω will be Z_1, the outward orientation variable. Appending an error term, u_{it}, equation (9) then becomes

$$\begin{aligned}\widehat{TFP}_{it} &= a_0 + \sum_{i=1}^{N-1} a_i D_i + \sum_{t=1}^{T-1} a_t D_t + \sum_{s=2}^{3} b_s Z_{sit} + \alpha \widehat{M}_{it} + \widehat{H}_{it}\theta(H_{it}, \Omega_{it}) + u_{it} \\ &= X_{it}\beta + \widehat{H}_{it}\theta(H_{it}, \Omega_{it}) + u_{it} \qquad (10)\end{aligned}$$

where $X_{it} = (D_i, D_t, Z_{sit}, \widehat{M}_{it})$ and the error term satisfies $E(u_{it}|X_{it}, H_{it}, \widehat{H}_{it}, \Omega_{it}) = 0$.

The central issue in (10) is the estimation of the $\theta(\cdot)$ function. The estimation approach adopted here is based on the smooth coefficient semiparametric model (see Fan, 1992; Fan and Zhang, 1999; Cai *et al.*, 2000a,b; Li *et al.*, 2002; Kourtellos, 2003). It is a generalization of varying coefficient models and follows the local polynomial linear regression of Stone (1977) and Fan (1992), as well as the widely used Nadaraya–Watson constant kernels.

The general description of the method follows. The data are given as $\{Y_i, W_i\}$, $i = 1, \ldots, n$, a realization from an i.i.d. random vector $\{Y, W\}$, where for notational simplicity we suppress the observation subscript *it* as $i = 1, \ldots, n$, with $n = N \times T$. The covariates are defined on $\mathcal{W} \subseteq \Re^{p+1}$. In our model we let Y_i denote *TFP*, and let $V_i = \{H_i, \Omega_i\}$, the dimension of the X-vector is l and that of the V-vector is q, $p = l + q$. Then we have $W_i = \{X_i, V_i, \widehat{H}_i\}$. We rewrite equation (10) as

$$Y_i = E(Y_i|W_i) + u_i = X_i\beta + \widehat{H}_i\theta(V_i) + u_i \qquad (11)$$

The presence of the linear part in equation (11) makes this model more general than the smooth coefficient model of Fan and Zhang (1999) (see Zhang *et al.*, 2002). We have followed the two-step approach due to Zhang *et al.* (2002). The coefficients of the linear part are estimated in the first step using polynomial fitting and an initial small bandwidth chosen by cross-validation (see Hoover *et al.*, 1998). These estimates are averaged to yield the first step linear part estimates. They are then used to redefine the dependent variable and return to the simple smooth coefficient environment of Fan and Zhang (1999) and Li *et al.* (2002) in equation (12):[6]

$$Y_i^* = \widehat{H}_i\theta(V_i) + u_i^* \qquad (12)$$

where Y^* and u^* denote the redefined dependent variable and error term.

The coefficient $\theta(v)$ evaluated at a particular value of V, say v, is a smooth but unknown function of v. In Appendix B we discuss how this function can take specific parametric formulations (such as linear) that can be tested against the general unknown specification. One can estimate $\theta(v)$ using a local least squares approach:

$$\begin{aligned}\widehat{\theta}(v) &= \left[(n\lambda^q)^{-1}\sum_{j=1}^{n}\widehat{H}_j^{\,2} K\left(\frac{v_j - v}{\lambda}\right)\right]^{-1}\left[(n\lambda^q)^{-1}\sum_{j=1}^{n}\widehat{H}_j Y_j^* K\left(\frac{v_j - v}{\lambda}\right)\right] \\ &= [B_n(v)]^{-1}C_n(v)\end{aligned}$$

[6] We have also followed an alternative two-step approach, where first we project off the influence of the linear part (the X) from the Y and the $\widehat{H}$ variables through least squares. In the second step, the residual projections are used to redefine the Y and $\widehat{H}$ variables and then used to estimate the unknown smooth coefficient functions. The coefficient function estimates from this method and the one by Zhang *et al.* (2002) are very similar.

where

$$B_n(v) = \frac{\sum_{j=1}^{n} \widehat{H}_j^2 K\left(\frac{v_j - v}{\lambda}\right)}{n\lambda^q}, \qquad C_n(v) = \frac{\sum_{j=1}^{n} \widehat{H}_j Y_j^* K\left(\frac{v_j - v}{\lambda}\right)}{n\lambda^q}$$

$K(\cdot)$ is a kernel function and λ is the smoothing parameter for sample size n.

The intuition behind the above local least squares estimator is straightforward. Let us assume that v is a scalar and $K(\cdot)$ is a uniform kernel. In this case the expression for $\widehat{\theta}(v)$ becomes

$$\widehat{\theta}(v) = \left(\sum_{|v_j - v| \leq \lambda} \widehat{H}_j^2\right)^{-1} \left(\sum_{|v_j - v| \leq \lambda} \widehat{H}_j Y_j^*\right)$$

where $\widehat{\theta}(v)$ is simply a least squares estimator obtained by regressing Y_j^* on $\widehat{H}_j$ using the observations of $(\widehat{H}_j, Y_j^*)$ that their corresponding v_j is close to v ($|v_j - v| \leq \lambda$). Since $\widehat{\theta}(v)$ is a smooth function of v, $|\theta(v_j) - \theta(v)|$ is small when $|v_j - v|$ is small. The condition that $n\lambda^q$ is large ensures that we have sufficient observations within the interval $|v_j - v| \leq \lambda$ when $\theta(v_j)$ is close to $\theta(v)$. Therefore, under the conditions that $\lambda \to 0$ and $n\lambda^q \to \infty$, one can show that the local least squares regression of Y_j^* on $\widehat{H}_j$ provides a consistent estimate of $\theta(v)$. In general it can be shown that

$$\sqrt{n\lambda^q}(\widehat{\theta}(v) - \theta(v)) \to N(0, \Phi)$$

where Φ can be consistently estimated. The estimate of Φ can be used to construct confidence bands for $\widehat{\theta}(v)$.

3.2. Data

The model is estimated with annual data for 51 countries during 1971–87. The main sources of data are: Nehru *et al.* (1995) for output, physical capital and human capital stock, the *National Accounts Statistics* of the United Nations (UN) for the share of employee compensation in national income, the Summers–Heston database for the number of workers, Freedom House (2000) for the indicators of political and civil freedom, and Sachs and Warner (1995) for the indicator of outward orientation. In order to arrive at an estimate of the labour share of national income, we have followed Gollin (2002) in adjusting the data on compensation of employees from the UN (which do not account for the self-employed) with data on the number of self-employed obtained from the *Year Book of Labour Statistics* of the International Labour Organization (ILO). As Gollin points out, this adjustment results in increasing the average estimate of the labour share of national income from 47.2% to 65.4% and reducing its standard deviation from 13.7% to 10.9%. The adjusted share of labour income in Gollin's sample (of 19 countries) ranges from a low of 48.5% to a high of 87.2%. Our estimates of the labour share of national income for 51 economies are in Table AI. They range from a low of 42.5% to a high of 74.6%, with an average share of labour income equal to 59.4% and a standard deviation equal to 9.9%. Though our estimates are somewhat lower than Gollin's, our estimate of both the average labour share and its standard deviation is not much different from Gollin's.[7]

[7] Appendix A contains a complete listing of the data sources and the method of construction of all the variables used in the estimation.

The factor constraining data availability to 51 countries and post-1970 data is the UN, *National Accounts Statistics*, which provide no information on employee compensation before 1970. The final year of our study (1987) is determined by the availability of estimates of human capital stock from Nehru *et al*. We have a total of 867 observations. We estimate the model with annual data rather than averaging the data over, for example, 5-year intervals in order to maximize sample size. It is well known that semiparametric estimation techniques require relatively large sample sizes. We fully recognize that resorting to annual data may introduce business cycle effects in estimation. As indicated in the previous subsection, all our estimated models include time-specific dummy variables that will capture, in part, such business cycle phenomena. In addition, as will be mentioned in the following subsection, we have experimented with averaging the data over longer subperiods (5-year periods), obtaining qualitatively similar results with averaged data. In the trade-off between averaged data (fewer business cycle effects) and annual data (more efficient semiparametric estimates), we have chosen to present the annual data results.

3.3. Empirical Results

We estimate several alternative specifications of (10). Before presenting the results, it is generally acknowledged that the measurement of the contribution of labour in national income accounts is beset with errors. In order to reduce the possibility that our results are driven by measurement error, we regressed the observed labour share on country characteristics such as the capital–labour ratio and per capita GDP. We then used estimates from this regression as our measure of the share of labour in national income. The results with the 'smoothed' labour share are very similar to those using the observed labour shares. In what follows, we report results based on 'smoothed' labour shares. Results based on measured labour shares are quite similar and are available on request.

Estimation results are presented in Table I. Column 1 of Table I shows estimates of the model in (10) assuming that θ is a constant, i.e., linear contribution for human capital. In this specification, the exogenous rate of technological change depends only on country- and time-specific fixed effects. Column 2 shows estimates of the linear model where the exogenous rate of technological change depends, in addition to country- and time-specific fixed effects, on outward orientation and civil and political freedoms. Column 4 of Table I shows the estimates of the linear component of the semiparametric counterpart to the model of column 2. Following Racine and Li (2004), we combine the standard Gaussian kernel for continuous variables with the uniform kernel for a discrete variable and choose the bandwidth by cross-validation. The nonparametric component of equation (12), the estimates of function θ, will be examined using graphical tools.

The estimate of scale (α) is insignificant in both linear and semiparametric models; this is true using either method for constructing the user cost of capital (see Appendix A). Therefore the null of constant returns to scale (in labour and capital) cannot be rejected. In what follows we present results on the assumption that payments to capital and labour exhaust total output; results using the other methods for constructing the user cost of capital are similar. Columns 1 and 2 also show that the coefficient estimate for human capital in the linear model is insignificant, as reported previously in the literature.

We have performed a number of robustness checks on our results. First, there are two widely recognized limitations that hamper estimation whether one attempts to estimate a TFP growth model as in (9) or a production function directly as in (5). The first concerns the aggregation of different types of labour and capital, an issue raised by Jorgenson and Griliches (1967) and recently

Table I. Parameter estimates (heteroskedastic std. errors)

Variable	Linear			Semiparametric	
	(1)	(2)	(3)	(4)	(5)
Trade (Z_1)		0.0133		0.0134	
		(0.0095)		(0.0095)	
Political (Z_2)		0.0006	0.0005	0.0006	0.0004
		(0.0023)	(0.0024)	(0.0024)	(0.0024)
Civil (Z_3)		0.0001	0.0003	0.0005	0.0009
		(0.0025)	(0.0025)	(0.0025)	(0.0025)
Scale ($\widehat{M}$)	0.1135	0.0808	0.0753	0.0939	0.0855
	(0.1446)	(0.1499)	(0.1495)	(0.1490)	(0.1476)
Human ($\widehat{H}$)	0.1874	0.1812	0.1060		
	(0.2374)	(0.2362)	(0.2815)		
Interaction ($\widehat{H} \times Z_1$)			0.2072		
			(0.3296)		
R^2	0.208	0.211	0.211	0.211	0.213

Note: Country- and time-specific dummies are included in all models.

by Barro (1999). Data limitations prevent us from pursuing this. The second is endogeneity. Estimating equation (10) may be problematic due to the possible endogeneity of $\widehat{H}$. This issue has been raised previously in the TFP growth literature (Griliches, 1973 contains a succinct discussion). The growth of human capital may respond to exogenous shocks to productivity growth (the $\widehat{A}$ term in (9)) and therefore the estimated coefficient on $\widehat{H}$ would proxy, in part, for exogenous technological progress.[8] By way of accounting for this endogeneity, we have instrumented for $\widehat{H}$. Finding suitable instruments for $\widehat{H}$ presents problems. We have used the dummy variables (D_i and D_t) as well as (twice) lagged values of (the logarithm of) the quantity and the price of labour as instruments. We have also experimented with other combinations of instruments (e.g. lagged values of the quantity of output or capital and the share of labour), arriving at similar results. In order to obtain an optimal instrument for $\widehat{H}$, the dummy variables are entered parametrically while the other three instruments are entered nonparametrically. The instrumental values of $\widehat{H}$ are then used in equation (10). The results are similar to those reported above.[9]

We have conducted a number of specification tests. First, we tested the null of a linear specification against the smooth coefficient specification of equation (10) using the test statistic described in Appendix B. The test statistic is bootstrapped, though we could reject the null hypothesis also using the nominal asymptotic critical values. The value of the test statistic is 2.16. We then tested the null of the smooth coefficient model in (10) against a partially linear semiparametric alternative where the nonparametric part contains $\widehat{H}$ and $V = \{H, Z_1\}$. The null is expressed as $E(Y_i|W_i) = X_i\beta + \theta(V_i)\widehat{H}_i$ and the alternative is $E(Y_i|W_i) = X_i\beta + g(V_i, \widehat{H}_i)$. In the alternative model, Z_1 interacts more generally with $\widehat{H}$ and H than in equation (10). We use a version of the test proposed by Fan and Li (1996) where we again bootstrapped the test statistic.

[8] Moreover, if human capital and the traditional inputs are correlated, the effect of human capital on TFP growth will be biased.

[9] In addition, and in order to check the consistency of our results, we have re-estimated the model averaging our data over three periods: 1971–76, 1977–82 and 1983–87. The use of such data may help in identifying more accurately the long-run impact of human capital by increasing the signal-to-noise ratio in the data. The disadvantage is that sample size is reduced substantially (from 867 to 153 observations), which will reduce the efficiency of the semiparametric estimates. The results (available on request) are qualitatively very similar to the complete sample.

The value of the test statistic is 0.20 and the bootstrap critical value is 1.11. We failed to reject the null (equation (10)) in that case. We also tested equation (10) against a general nonparametric model using again a bootstrap version of the test by Fan and Li (1996). In that case the null hypothesis of the semiparametric specification can be expressed as $E(Y_i|W_i) = X_i\beta + \theta(V_i)\widehat{H}_i$ and the alternative hypothesis as $E(Y_i|W_i) = f(W_i) \neq X_i\beta + \theta(V_i)\widehat{H}_i$. Under the null hypothesis, this nonparametric functional form (*NPFF*) test statistic follows the asymptotic standard normal distribution. The value of the Fan–Li *NPFF* statistic is 0.63 with a bootstrap critical value of 1.14. Therefore, the null hypothesis of a semiparametric specification cannot be rejected against the alternative.

The other specifications in Table I (columns 3 and 5) introduce the role of trade orientation (Z_1) as a determinant of the contribution of human capital to economic growth. The specification in column 5 assumes that $\theta(\cdot)$ (or, as will be argued shortly, the elasticity of human capital) depends on the level of human capital and trade orientation and the specification in column 3 is the linear approximation counterpart to this model by introducing an interaction term between human capital and trade orientation. That the elasticity (and consequently the rate of return) of human capital should depend on the level of human capital is straightforward. What is less straightforward and requires explanation is why the output elasticity of human capital should depend on a country's trade orientation. Pissarides (1997) claims that the return to human capital increases relative to the return to unskilled labour (as manifested by a widening gap in wages between skilled and unskilled labour) as developing countries liberalize their trade regime. This is due to the transfer of skill-biased technology to developing economies from developed countries as these economies increase their exposure to the world economy. All the economies that have moved towards increased trade openness during our sample period are low- or middle-income. We model this effect for the linear model in column 3 by introducing an interaction term between human capital and trade orientation and for the semiparametric model in column 5 by estimating the specification in (10) where the $\theta(\cdot)$ function depends on both human capital and outward orientation or $\theta(H_{it}, Z_{1it})$.

The results in Table I show no evidence that greater political or civil freedoms promote TFP growth in either linear or semiparametric models, but there is marginal evidence of the beneficial effect of outward orientation on the growth of productivity.

3.4. The Output Elasticity of Human Capital

Of key importance to our work is the derivative of TFP growth with respect to human capital growth or $\theta(\cdot)$. It is important to note that this derivative corresponds to the output elasticity of mean years of schooling (H), i.e.,

$$\varepsilon_{YH_{it}} = \frac{df}{dH_{it}}\frac{H_{it}}{Y_{it}} = \varepsilon_{E_{it}}\eta_{H_{it}} + \varepsilon_{H_{it}} = \theta(\cdot) \tag{13}$$

Figure 1 plots pointwise estimates of the output elasticity, $\theta(\cdot)$, on the vertical axis and the level of human capital, H_{it}, on the horizontal axis, along with the 95% confidence bands for the specification in column 4 of Table I. Inspection of Figure 1 shows that the relationship between elasticity and human capital is nonlinear. In addition, apart from a small range, the 95% confidence bands do not include zero. The average output elasticity for each country, along with its standard error, are in Table II, corresponding to the specifications in columns 4 and 5 of Table I. The average elasticities of specification 4 lie within a relatively wide but 'sensible' range: 0.07 to

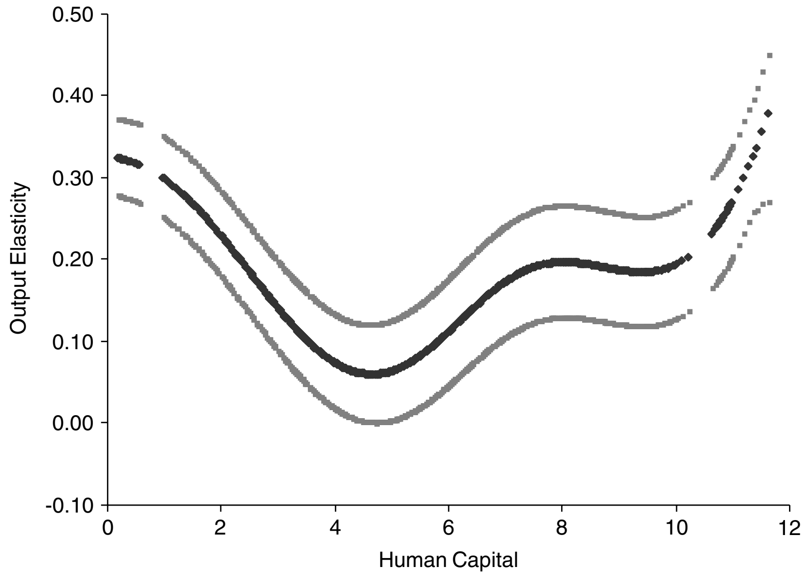

Figure 1. Human capital output elasticity, $\widehat{\theta}(H)$ (Model 4)

0.28.[10] There is a large variation in the estimates of elasticity across countries, casting doubt on the widespread assumption in the literature of uniform elasticity across countries. The elasticity estimates are largest for the high-income economies: the (unweighted) average elasticity for the 19 highest income economies in our sample is 0.19 (and ranges from a low of 0.13 for Spain to a high of 0.28 for the United States). Figure 1 indicates that for economies with relatively low levels of human capital there are decreasing returns to scale for human capital. For middle-level human capital economies, returns to human capital are increasing while for the highest human-capital economies, there are constant or mildly increasing returns to human capital.

In many developing economies, our estimate of the output elasticity is low. The average of the estimate of the elasticity for the countries in Table II identified by the World Bank (1989) as low- or middle-income in 1987 is 0.13 (the average elasticity of the low-income economies is 0.16 and for medium-income is 0.11). It may be that structural obstacles prevent the efficient usage of human capital in developing countries. It is well known that human capital requires complementary 'advanced' technologies, a factor in scarce supply in developing economies but a factor that can be made available through increased exposure to international trade. The lack of complementary technologies implies that increased levels of education are frequently directed towards socially-wasteful or directly-unproductive activities or may come up against greater incidence of unemployment, a point noted by Pritchett (2001) and Krueger and Lindahl (2001). There is certainly a large body of anecdotal evidence describing university graduates in developing countries devoting their talents to rent-seeking activities or frequently becoming unemployed (or

[10] We exclude from the discussion and Figures 1 and 2 Ethiopia and Ireland, the countries with the highest and lowest average values for human capital and also the highest and lowest rates of growth of human capital (see Table AI). Semiparametric estimates at the 'tails' of the distribution tend to be the least accurate.

Table II. Output elasticities and rates of return (average by country, std. error in parentheses)

Country	Output elasticity		Rate of return	
	(4)	(5)	(4)	(5)
Algeria	0.148	0.106	0.135	0.102
(DZA)	(0.051)	(0.080)	(0.062)	(0.082)
Australia	0.165	0.279	0.056	0.095
(AUS)	(0.022)	(0.053)	(0.006)	(0.019)
Austria	0.192	0.067	0.047	0.016
(AUT)	(0.0008)	(0.005)	(0.001)	(0.002)
Belgium	0.196	0.140	0.063	0.045
(BEL)	(0.0004)	(0.021)	(0.0005)	(0.006)
Canada	0.188	0.040	0.046	0.010
(CAN)	(0.003)	(0.016)	(0.0007)	(0.004)
Colombia	0.086	0.052	0.060	0.034
(COL)	(0.024)	(0.152)	(0.020)	(0.097)
Costa Rica	0.151	−0.040	0.072	−0.020
(CRI)	(0.037)	(0.126)	(0.016)	(0.060)
Denmark	0.192	0.074	0.050	0.020
(DNK)	(0.003)	(0.025)	(0.001)	(0.007)
Ecuador	0.072	0.293	0.038	0.158
(ECU)	(0.011)	(0.254)	(0.005)	(0.134)
El Salvador	0.082	−0.023	0.047	−0.011
(SLV)	(0.021)	(0.066)	(0.015)	(0.037)
Ethiopia	0.321	0.321	1.934	1.934
(ETH)	(0.003)	(0.003)	(0.675)	(0.675)
Finland	0.190	0.060	0.056	0.018
(FIN)	(0.005)	(0.040)	(0.002)	(0.012)
France	0.196	0.131	0.061	0.040
(FRA)	(0.0008)	(0.026)	(0.0003)	(0.008)
Germany	0.196	0.110	0.064	0.036
(DEU)	(0.0008)	(0.013)	(0.0009)	(0.004)
Greece	0.195	0.112	0.044	0.025
(GRC)	(0.002)	(0.033)	(0.0004)	(0.008)
Iceland	0.191	0.127	0.062	0.042
(ISL)	(0.005)	(0.072)	(0.002)	(0.024)
India	0.158	0.126	0.097	0.079
(IND)	(0.039)	(0.056)	(0.036)	(0.044)
Indonesia	0.109	0.432	0.085	0.331
(IDN)	(0.037)	(0.001)	(0.033)	(0.014)
Ireland	1.263	1.249	0.388	0.383
(IRL)	(0.445)	(0.473)	(0.131)	(0.140)
Italy	0.181	0.237	0.046	0.061
(ITA)	(0.011)	(0.042)	(0.002)	(0.012)
Jamaica	0.186	0.155	0.293	0.245
(JAM)	(0.009)	(0.117)	(0.020)	(0.184)
Japan	0.257	0.117	0.069	0.032
(JPN)	(0.013)	(0.017)	(0.005)	(0.005)
Kenya	0.145	0.103	0.099	0.073
(KEN)	(0.048)	(0.078)	(0.043)	(0.058)
Korea	0.122	0.342	0.060	0.169
(KOR)	(0.050)	(0.085)	(0.024)	(0.044)
Madagascar	0.168	0.140	0.127	0.108
(MDG)	(0.039)	(0.055)	(0.040)	(0.049)
Malawi	0.114	0.062	0.073	0.040
(MWI)	(0.004)	(0.007)	(0.006)	(0.006)

Table II. (*Continued*)

Country	Output elasticity (4)	Output elasticity (5)	Rate of return (4)	Rate of return (5)
Malaysia	0.079	0.413	0.038	0.203
(MYS)	(0.020)	(0.024)	(0.008)	(0.022)
Mauritius	0.110	0.374	0.068	0.229
(MUS)	(0.027)	(0.033)	(0.018)	(0.016)
Mexico	0.073	−0.067	0.044	−0.040
(MEX)	(0.014)	(0.124)	(0.008)	(0.072)
Netherlands	0.197	0.122	0.061	0.038
(NLD)	(0.0006)	(0.014)	(0.0006)	(0.004)
New Zealand	0.185	0.202	0.077	0.084
(NZL)	(0.014)	(0.249)	(0.007)	(0.102)
Norway	0.188	0.043	0.056	0.013
(NOR)	(0.002)	(0.014)	(0.0007)	(0.004)
Pakistan	0.252	0.247	0.242	0.237
(PAK)	(0.019)	(0.021)	(0.043)	(0.045)
Panama	0.129	−0.092	0.068	−0.048
(PAN)	(0.041)	(0.085)	(0.022)	(0.045)
Paraguay	0.087	−0.153	0.045	−0.077
(PRY)	(0.006)	(0.002)	(0.004)	(0.002)
Philippines	0.150	−0.045	0.075	−0.022
(PHI)	(0.036)	(0.120)	(0.019)	(0.060)
Portugal	0.067	0.423	0.029	0.184
(PRT)	(0.010)	(0.011)	(0.004)	(0.007)
Sierra Leone	0.264	0.260	0.328	0.324
(SLE)	(0.028)	(0.031)	(0.101)	(0.104)
Singapore	0.110	0.366	0.053	0.181
(SGP)	(0.046)	(0.069)	(0.019)	(0.044)
Spain	0.133	0.340	0.065	0.166
(ESP)	(0.028)	(0.046)	(0.013)	(0.023)
Sri Lanka	0.092	0.047	0.039	0.020
(LKA)	(0.020)	(0.264)	(0.010)	(0.113)
Sweden	0.187	0.033	0.039	0.007
(SWE)	(0.002)	(0.009)	(0.0003)	(0.002)
Switzerland	0.145	0.325	0.053	0.119
(CHE)	(0.014)	(0.025)	(0.005)	(0.008)
Tanzania	0.257	0.252	0.310	0.305
(TZA)	(0.033)	(0.038)	(0.106)	(0.110)
Thailand	0.070	0.421	0.035	0.213
(THA)	(0.011)	(0.012)	(0.005)	(0.008)
Turkey	0.099	0.021	0.076	0.020
(TUR)	(0.034)	(0.070)	(0.033)	(0.055)
United Kingdom	0.189	0.028	0.052	0.008
(GBR)	(0.005)	(0.006)	(0.002)	(0.002)
United States	0.284	0.152	0.101	0.054
(USA)	(0.043)	(0.056)	(0.019)	(0.022)
Venezuela	0.087	−0.104	0.040	−0.047
(VEN)	(0.029)	(0.049)	(0.012)	(0.022)
Zambia	0.132	0.074	0.099	0.061
(ZMB)	(0.058)	(0.101)	(0.056)	(0.080)
Zimbabwe	0.076	−0.026	0.045	−0.015
(ZWE)	(0.016)	(0.043)	(0.011)	(0.025)

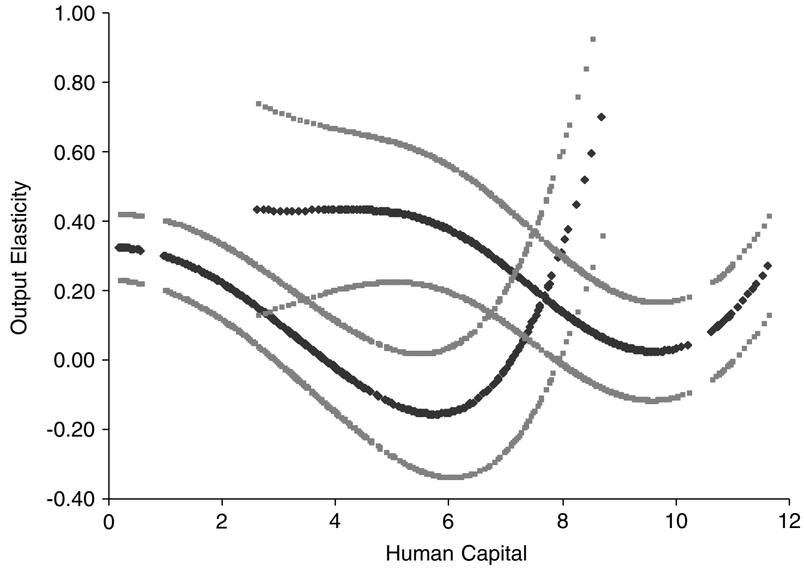

Figure 2. Human capital output elasticity, $\widehat{\theta}(H, Z_1)$ (Model 5)

underemployed in work that is not commensurate with their level of education) upon graduation, and being forced to enter the unofficial economy.

Figure 2 shows pointwise estimates of the elasticity of human capital for the model in column 5 of Table I, i.e., when we assume that a country's outward orientation (as well as its level of human capital) is a determinant of the human capital elasticity. The figure also shows the 95% confidence bands. One additional interesting conclusion emerges from Figure 2. Two distinct nonlinear relationships can be observed: the upper of the two curves (and the confidence bands) corresponds to economies that are open ($Z_1 = 1$) while the lower (and the confidence bands) corresponds to economies that are closed ($Z_1 = 0$). Except for a very small range, there is very little overlap between the two sets of confidence bands. It is clear that, *ceteris paribus*, the elasticity of human capital is higher the more open an economy. To the best of our knowledge, this is the first systematic empirical confirmation of the oft cited proposition that greater exposure to international trade brings access to foreign technology that complements human capital and raises its marginal contribution to aggregate productivity.

3.5. The Social Rate of Return to Human Capital

As mentioned in the introduction, numerous studies have estimated the return to human capital (education) on the basis of micro survey data. Psacharopoulos (1994) summarizes the literature and provides indicative rates of return for various countries. On the other hand, the literature provides no studies of estimates of the return to human capital based on aggregate (macro) data for a wide cross-section of economies. Krueger and Lindahl (2001) and Topel (1999) attempt to link the micro studies with the cross-country growth accounting literature, but their methodology does not allow country-specific estimates of the return to human capital.

The methodology presented in this paper allows us to retrieve an estimate of the social return to human capital that is country (and time)-specific. This is equal to the marginal benefit over the marginal cost of an additional unit of human capital. The marginal benefit is defined as the additional units of output per worker gained (or lost) as a result of a unitary increase in human capital (mean years of schooling) or, using (13), $\varepsilon_{YH}\ (Y/L)/H$. The marginal cost of an additional unit of human capital is assumed to equal the real wage rate for unskilled or 'raw' labour, P_L^b/P_Y, the opportunity cost of schooling. Computing the wage rate for raw labour (P_L^b) is problematic because there are no widely available estimates that are comparable across countries. In Appendix A we present a computational method that is consistent with estimated rates of return to education from micro studies.

Bils and Klenow (2000) have assembled representative returns (in terms of higher wages) to an additional year of schooling from a variety of sources. These returns are from estimates of Mincerian wage functions with micro data; this is the private return to education. They are available for 33 of the 51 countries in our sample. These range from a low of 2.6% for Sweden to a high of 28% for Jamaica. As Krueger and Lindahl (2001) point out, the social return to education can be higher or lower than the private return. The former is likely if higher levels of human capital engender technological progress not captured by the private return or reduction in social variables (e.g. crime or fertility rates) or 'more informed political decisions'. The latter is likely if education serves to raise social status without raising productivity.

Table II reports our average estimate of the social return to human capital for each country along with its standard error for the specifications in columns 4 and 5 of Table I. For the 33 countries for which private rates of return are available, we note that for the 12 high-income economies the social rate of return to human capital is approximately equal to, or in many cases exceeds, the private rate of return. The same is true for four of the five low-income economies. On the other hand, for 14 of the 16 middle-income economies, the social rate of return is lower than the private rate of return. It appears that positive externalities to human capital are concentrated either at low levels of development, where the benefits of education in terms of lower fertility, opportunities for enhanced nutrition and sanitation manifest themselves, or at high levels of development, where the benefits of technological spillovers or 'more informed political decisions' come into play. It is also interesting that, for the model assuming elasticity depends on outward orientation, the estimates of the social rate of return to human capital for all but one of the six South East Asian economies (Indonesia, Korea, Malaysia, Singapore and Thailand) are among the highest of all economies. These countries have been praised incessantly for their outward orientation and for their efforts in providing widespread access to educational opportunities for their population. The sole exception is the Philippines, a country notorious among the South East Asian economies for rent-seeking and other unproductive activities.[11] Other outward oriented economies (e.g. Mauritius) also have high and significant estimates of the social rate of return to human capital.

In conclusion, we find that the social return to human capital is highest for low-income economies and for the richest economies with the middle-income economies lagging behind these two groups. As pointed out by many researchers (e.g. Psacharopoulos, 1994), private returns to education are higher for developing economies and lower for rich economies. Putting these

[11] Indonesia and the Philippines are ranked as the two most corrupt economies of those in our sample according to the Corruption Perceptions Index of Transparency International (2001) during the 1980–85 period. The Indonesian economy, however, is classified as open by the Sachs/Warner criterion while the Philippines is considered to be closed for all the years of our sample, a factor that helps explain the difference in estimates.

together, social rates of return to human capital appear to be in line with private rates at the two ends of the development continuum and lower than private rates for the middle-income economies. Taking a sweeping overview of our results, we might hazard a very tentative conclusion: the positive externalities to human capital accumulation emphasized by a number of theories of endogenous growth exist mainly for low- and high-income economies, while there is no evidence of such externalities for the middle-income group.

4. CONCLUSION

This paper looks at the impact of human capital accumulation on the growth of output. It argues that previous studies have been unable to discern a significant effect because they assume that the contribution of the 'traditional' inputs (labour and physical capital) as well as human capital is constant both across countries and time. We examine both assumptions. First, we use data on the contribution of labour and physical capital that vary across countries and time to remove the effects of the growth of traditional inputs from output growth. Second, we show that the (growth of the) resulting TFP index depends on the growth of human capital. Importantly, we allow the contribution of human capital to TFP growth to vary across countries and time by modelling this contribution by a general function that we estimate via semiparametric techniques.

Our analysis yields several conclusions. First, our semiparametric methodology enables us to recover estimates of the elasticity of human capital with respect to output that differ across countries and time. The average output elasticity varies substantially across countries. In some cases the estimate is negligible, providing empirical support to the hitherto unexplored proposition that human capital accumulation might not yield a beneficial effect on output. We also find that, *ceteris paribus*, the human capital elasticity of aggregate output is higher for outward-looking economies. Finally, we compute social rates of return to human capital from aggregate data. There is a wide dispersion of estimates. Our study is an important first step in deriving elasticities and rates of return to human capital based on aggregate (national) data in contrast to the preponderance of private rates of return from Mincerian earnings functions.

APPENDIX A

Our sample consists of 51 countries during 1971–87 or a total of 867 observations. All variables are calculated for each country and time period so subscripts are omitted.

Output: Output in constant (1987) domestic prices (Y_C) is from the Nehru *et al.* (1995) database; the price of output (q_Y) is the GDP deflator (from the *World Tables* of the World Bank). Output in constant 1987 US dollars (Y) is Y_C divided by the base year PPP (PPP_Y^0, from the Summers–Heston database). The corresponding price index is defined as $P_Y = q_Y PPP_Y^0$.

Output share of labour: We compute the share of labour in GDP (s_{YL}) first by collecting data on the compensation of employees paid by resident producers (as a percentage of GDP) from various issues of the *National Accounts Statistics* of the United Nations (table I.3). This data, however, does not account for self-employment, a fact also pointed out by Gollin (2002). Therefore, we also obtain data on the number of self-employed (employers and own account workers) in each country as a proportion of the number of employees and use these to adjust the UN data accordingly. This adjustment assumes implicitly that the average wage of employees and self-employed workers is the same, an assumption also made by Gollin (2002). Data on the number of employees and self-employed is from various issues of the *Year Book of Labour Statistics* of the International Labour Organization (ILO). Because complete data on the number of employees and self-employed is not available for all countries, some interpolation was necessary.

Labour: We estimate total labour compensation by multiplying s_{YL} by $q_Y Y_C$. We construct a price index of labour (q_L) by dividing labour compensation by the number of workers ($\aleph$, from the Summers–Heston database) and normalizing this number to equal 1 in 1987 (base year). Labour quantity in constant domestic prices (L_C) is constructed implicitly by dividing labour compensation by q_L. Labour in constant 1987 US dollars (L) is L_C divided by the base year PPP for labour (PPP_L^0). The corresponding labour price index is defined as $P_L = q_L PPP_L^0$. The PPP for labour of country i is constructed by dividing the labour cost per worker at the base year by the corresponding value in the USA, i.e., $PPP_L^0 = (q_L^0 L_C^0/\aleph^0)^i/(q_L^0 L_C^0/\aleph^0)^{\mathrm{US}}$.

Physical capital: Physical capital in constant (1987) domestic prices (K_C) is from the Nehru *et al.* (1995) database; it is built from investment series (from the World Bank database) via the perpetual inventory method. The acquisition price of investment (q_I) is the investment deflator (from the *World Tables* of the World Bank). There are no cross-country data on the rental price of capital. Therefore, we constructed these in three ways. First, we assume that the value of labour and capital exhausts total output and the rental price of capital is $q_K = (q_Y Y_C - q_L L_C)/K_C$. Second, $q_K = q_I[r - (\widehat{q}_I) + \delta]$, where r is the rate of return to capital, $(\widehat{q}_I)$ is capital gains and δ is the depreciation rate. We assume the real return to capital $r - (\widehat{q}_I)$ differs between low-, middle- and high-income countries. Following representative estimates reported by Harberger (1998), we set this rate equal to 9% for low-income, 5% for middle-income and 2% for high-income countries. The rate of depreciation is that used by Nehru *et al.* (1995); it is 4% for all countries. Finally, we assume a constant real return to capital of 6%. This is close to the average for all the countries in Harberger (1998) and also to the average real long-term government bond yield for the countries with available data in our sample (data on long-term government bond yields is from the *International Financial Statistics* of the International Monetary Fund). Physical capital in constant 1987 US dollars (K) is K_C divided by the base year PPP for investment (PPP_I^0, from the Summers–Heston database). The corresponding capital price index is defined as $P_K = q_K PPP_I^0$. Finally, we define the cost-share of capital $S_K = P_K K/C$ and labour $S_L = P_L L/C$ where $C = P_K K + P_L L$.

Human capital: Our measure of the human capital stock is the total (primary, secondary and tertiary) number of years of schooling in the working age population; estimates are from the Nehru *et al.* (1995) database.

Wage rate for raw labour: We decomposed the share of labour in GDP (s_{YL}) into a component due to raw labour and one due to human capital. We accomplished this by assuming that human capital raises the (unknown) wage of raw labour (P_L^b) for country i as follows: $P_L^s = P_L^b e^{\rho H_s}$, where ρ is the return to schooling for country i and H_s is human capital (mean years of schooling) of country i for schooling level s. We consider four schooling levels: no schooling, primary, secondary and tertiary. The private rate of return to schooling for each country is from Bils and Klenow (2000). For countries where Bils and Klenow provide no data, we use the average rate of return for different regions (Sub-Saharan Africa, Asia, Latin America, North Africa/Middle East and OECD) from Psacharopoulos (1994). Human capital estimates by level of education for each country (H_s) are from the Nehru *et al.* (1995) database. The human capital component in s_{YL} (as a percentage of s_{YL}) for country i is then obtained as $\sum_s \ell^s(P_L^s - P_L^b)/\sum_s \ell^s P_L^s$, where ℓ^s is country i's share of the working age population with level of education s. These shares (ℓ^s) are from the Barro–Lee

Table AI. Data averages by country (%, 1971–87)

Country	Output share of labour (s_{YL})	Growth of				
		Output ($\widehat{Y}$)	Labour ($\widehat{L}$)	Physical capital ($\widehat{K}$)	Human capital ($\widehat{H}$)	Total factor productivity ($\widehat{TFP}$)
Algeria	44.19	4.60	3.35	7.29	4.05	−1.01
Australia	62.32	3.13	2.14	3.98	1.23	0.30
Austria	62.22	2.67	0.81	4.58	0.01	0.44
Belgium	64.73	2.31	0.74	3.01	0.43	0.76
Canada	60.26	3.89	2.30	4.63	0.92	0.66
Colombia	49.53	4.47	2.55	4.71	2.93	0.83
Costa Rica	54.83	3.89	3.56	6.29	2.02	−0.87
Denmark	61.98	2.30	0.98	3.39	0.67	0.40
Ecuador	43.03	5.71	2.71	5.60	2.54	1.44
El Salvador	49.03	1.54	1.90	4.69	2.75	−1.76
Ethiopia	64.61	2.51	2.17	4.91	7.42	−0.78
Finland	63.06	3.24	0.80	3.82	1.35	1.34
France	64.07	2.63	0.91	4.30	0.53	0.48
Germany	61.48	2.14	0.28	3.14	0.05	0.74
Greece	64.99	3.15	0.66	4.93	0.84	0.91
Iceland	59.25	5.12	2.27	4.79	1.52	1.83
India	69.58	3.82	1.99	4.51	3.33	1.07
Indonesia	73.22	6.15	2.23	10.41	3.30	1.70
Ireland	69.54	3.66	1.09	5.06	−1.21	1.38
Italy	65.44	2.97	0.52	3.62	0.98	1.41
Jamaica	72.76	−0.16	2.59	1.17	0.95	−2.29
Japan	65.06	4.06	0.87	7.37	0.24	0.86
Kenya	51.72	6.09	3.68	3.41	4.10	2.57
Korea	55.10	8.58	2.37	12.05	3.19	1.82
Madagascar	61.16	0.27	2.05	1.90	3.14	−1.74
Malawi	51.13	4.24	2.56	6.21	0.01	0.00
Malaysia	51.87	6.31	3.45	9.91	2.82	−0.22
Mauritius	48.92	5.51	2.25	3.15	1.80	2.80
Mexico	49.03	4.20	3.26	6.64	2.50	−0.70
Netherlands	62.98	2.15	1.41	3.21	0.27	0.08
New Zealand	60.47	2.25	1.71	3.26	1.77	−0.08
Norway	58.97	4.05	1.55	4.16	0.55	1.46
Pakistan	63.40	5.31	3.17	5.31	2.48	1.33
Panama	74.39	4.33	2.62	6.36	2.14	1.08
Paraguay	62.48	5.68	3.17	9.46	0.34	0.24
Philippines	58.47	3.43	2.49	6.31	1.87	−0.63
Portugal	63.34	3.52	1.51	4.79	1.54	0.72
Sierra Leone	45.63	1.39	1.57	2.38	5.23	−0.56
Singapore	43.90	7.42	3.08	12.37	3.64	−0.81
Spain	64.21	2.91	0.89	4.76	1.31	0.68
Sri Lanka	59.35	4.42	1.87	7.30	1.45	0.34
Sweden	66.23	1.96	0.85	2.95	0.67	0.40
Switzerland	66.42	1.42	0.55	3.52	0.59	−0.16
Tanzania	62.21	2.73	2.33	4.26	5.83	−0.02
Thailand	53.94	6.24	2.75	8.33	1.63	0.97
Turkey	42.53	5.19	2.06	6.35	3.09	0.84
UK	63.97	2.18	0.52	2.98	0.46	0.79
USA	65.78	2.80	1.86	2.89	0.47	0.59
Venezuela	52.25	1.87	4.08	4.55	3.44	−2.42
Zambia	62.87	1.21	3.03	0.54	4.94	−0.88
Zimbabwe	74.61	2.90	3.08	3.69	1.98	−0.40
Average	59.68	3.62	2.02	5.08	1.97	0.35
Std. Dev.	(9.86)	(3.78)	(1.22)	(3.25)	(1.75)	(3.31)

Note: See text for the calculation of total factor productivity growth, $\widehat{TFP}$.

database. Pritchett (1996) uses a similar procedure for identifying the human capital component of the total wage bill. Knowledge of the raw labour component of s_{YL} and the workforce is sufficient to calculate the wage rate for raw labour.

Political and civil freedoms: We measure political and civil freedoms according to the index compiled by Freedom House (2000). The range of the index is from 1 to 7, with higher values indicating lower degrees of freedom. Freedom House defines political rights as those that 'enable people to participate freely in the political process, which is the system by which the polity chooses authoritative policy makers and attempts to make binding decisions affecting the national, regional, or local community. In a free society, this represents the right of all adults to vote and compete for public office, and for elected representatives to have a decisive vote on public policies'. They define civil liberties to 'include the freedoms to develop views, institutions, and personal autonomy apart from the state'. The Freedom House indexes are frequently used in empirical research (see Rodrik, 2000).

Outward orientation: As Edwards (1998) discusses, measuring outward orientation is problematic. While a number of indices are available cross-sectional for a limited number of countries, there is only one index available on a consistent basis for a cross-section of countries across time. The Sachs and Warner (1995) binary index classifies a country as open (index equals 1) according to four criteria: the extent of the black market premium, distortions created by export marketing boards, the ideological nature of the regime (socialist or otherwise), and quota coverage on imports of intermediate and capital goods. The other commonly used measure of trade openness that is available across countries and time, the ratio of exports (or imports) to GDP, is a consequence of trade orientation rather than an independent indicator of openness. Moreover, the use of a binary indicator is in fact desirable since it allows us to present graphically differences in elasticities and rates of return between closed and open economies

Nehru *et al*. (1995) provide annual estimates of years of schooling for 1960–87 for 83 countries. The UN *National Accounts Statistics* do not provide consistent estimates of the compensation of employees (as a percentage of GDP) before 1970. In addition, the UN provides data for only 51 countries in the Nehru *et al*. database. These constraints limit our sample size to 51 countries during 1970–87, or a total of 867 observations of annual growth rates. Nehru *et al*. (1995) build their estimate of the stock of human capital from enrollment data using the perpetual inventory method. Thus, starting our sample in 1970 (instead of 1960) may provide a more accurate estimate. Moreover, most of the countries omitted from the Nehru *et al*. database are ranked (by them) in categories 3 or 4, indicating that there were substantial gaps in enrollment data and these had to be extrapolated.

APPENDIX B

We present a test statistic used by Li *et al*. (2002) to test a general linear model against the smooth coefficient semiparametric model. In our implementation we will use a bootstrap version of this test. Let y_i denote the dependent variable, and let x_i be $p \times 1$ and z_i be $q \times 1$ vectors of exogenous variables. Consider the following parametric model:

$$y_i = \alpha_0(z_i) + x_i^T \beta_0(z_i) + \varepsilon_i = (1, x_i^T) \begin{pmatrix} \alpha_0(z_i) \\ \beta_0(z_i) \end{pmatrix} + \varepsilon_i = X_i^T \delta_0(z_i) + \varepsilon_i \quad \text{(B1)}$$

where $\delta_0(z_i) = (\alpha_0(z_i), \beta_0(z_i)^T)^T$ is a smooth known function of z and $X_i^T = (1, x_i^T)$. For example, we could have $\alpha_0(z_i) = \alpha + z_i\theta$ and $\beta_0(z_i) = \beta$. Similarly, equation (B1) could allow for simple interactions of the x with z, where $\alpha_0(z_i) = \alpha + z_i\theta$ and $\beta_0(z_i) = \beta_1 + \beta_2 z$.

The alternative semiparametric model is when $\delta(z_i)$ is unknown and it is given by:

$$y_i = X_i^T \delta(z_i) + \varepsilon_i \tag{B2}$$

We can test the adequacy of (B1), the null hypothesis (H_0), against the semiparametric alternative (H_1) given by (B2) using the following test statistic:

$$\begin{aligned}\widehat{I}_n &= \frac{1}{n^2\lambda^q}\sum_i \sum_{j\neq i} X_i^T(y_i - X_i^T\widehat{\delta}_0(z_i))X_j(y_j - X_j^T\widehat{\delta}_0(z_j))K\left(\frac{z_j - z_i}{\lambda}\right) \\ &= \frac{1}{n^2\lambda^q}\sum_i \sum_{j\neq i} X_i^T X_j \widehat{\varepsilon}_i \widehat{\varepsilon}_j K\left(\frac{z_j - z_i}{\lambda}\right)\end{aligned}$$

where $\widehat{\varepsilon}_i$ denotes the residual from parametric estimation (under H_0). It can be shown that under H_0, $J_n = n\lambda^{q/2}\widehat{I}_n/\widehat{\sigma}_0 \to N(0, 1)$, where $\widehat{\sigma}_0$ is a consistent estimator of the variance of $\widehat{I}_n$. It can be shown that the test statistic is a consistent test for testing H_0 (equation (B1)) against H_1 (equation (B2)). We use a bootstrap version of the above test statistic since bootstrapping improves the size performance of kernel-based tests for functional form (see Zheng, 1996; Li and Wang, 1998).

ACKNOWLEDGEMENTS

The authors would like to thank Lant Pritchett and participants at seminars at the 14th European Economic Association meetings for constructive comments. Elena Ketteni provided excellent research assistance. The paper was initiated while the last two authors were visiting the University of Cyprus.

REFERENCES

Azariadis C, Drazen A. 1990. Threshold externalities in economic development. *Quarterly Journal of Economics* **105**: 501–526.
Barro R. 1991. Economic growth in a cross section of countries. *Quarterly Journal of Economics* **106**: 407–433.
Barro R. 1999. Notes on growth accounting. *Journal of Economic Growth* **4**: 119–137.
Barro R, Sala-i-Martin X. 1995. *Economic Growth*. McGraw-Hill: New York.
Benhabib J, Spiegel M. 1994. The role of human capital in economic development: evidence from aggregate cross-country data. *Journal of Monetary Economics* **34**: 143–174.
Bils M, Klenow PJ. 2000. Does schooling cause growth?" *American Economic Review* **90**: 1160–1183.
Cai Z, Fan J, Li R. 2000a. Efficient estimation and inferences for varying-coefficient models. *Journal of the American Statistical Association* **95**: 888–902.
Cai Z, Fan J, Yao Q. 2000b. Functional coefficient regression models for nonlinear time series models. *Journal of the American Statistical Association* **95**: 941–956.
Diewert WE. 1976. Exact and superlative index numbers. *Journal of Econometrics* **4**: 115–146.
Durlauf SN. 1993. Nonergodic economic growth. *Review of Economic Studies* **60**: 349–366.
Durlauf SN, Johnson P. 1995. Multiple regimes and cross country growth behavior. *Journal of Applied Econometrics* **10**: 365–384.

Durlauf SN, Kourtellos A, Minkin A. 2001. The local Solow growth model. *European Economic Review* **15**: 928–940.

Edwards S. 1998. Openness, productivity and growth: what do we really know? *The Economic Journal* **108**: 383–398.

Fan J. 1992. Design-adaptive nonparametric regression. *Journal of the American Statistical Association* **87**: 998–1004.

Fan Y, Li Q. 1996. Consistent model specification tests: omitted variables, parametric and semiparametric functional forms. *Econometrica* **64**: 865–890.

Fan J, Zhang W. 1999. Statistical estimation in varying-coefficient models. *The Annals of Statistics* **27**: 1491–1518.

Freedom House. 2000. Annual Survey of Freedom Country Scores 1972–73 to 1998–99. Available at http://www.freedomhouse.org/ratings (2000, January).

Gollin D. 2002. Getting income shares right. *Journal of Political Economy* **110**: 458–474.

Griliches Z. 1973. Research expenditures and growth accounting. In *Science and Technology in Economic Growth*, Williams BR (ed.). Macmillan: New York.

Hall R, Jones CI. 1999. Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* **114**: 83–116.

Harberger AC. 1998. A vision of the growth process. *American Economic Review* **88**: 1–32.

Hoover DR, Rice C, Wu CO, Yang LP. 1998. Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrica* **85**: 809–822.

International Labour Organization. *Yearbook of Labour Statistics*. International Labour Office: Geneva (various issues).

Jorgenson DW, Griliches Z. 1967. The explanation of productivity change. *Review of Economic Studies* **34**: 249–280.

Kalaitzidakis P, Mamuneas T, Savvides A, Stengos T. 2001. Measures of human capital and nonlinearities in economic growth. *Journal of Economic Growth* **6**: 229–254.

Klenow P, Rodríguez-Clare A. 1997. The neoclassical revival in growth economics: has it gone too far? In *NBER Macroeconomics Annual 1997*, Bernanke BS, Rotemberg JJ (eds). MIT Press: Cambridge, MA; 73–103.

Kourtellos A. 2003. Modeling parameter heterogeneity in cross-country growth regression models. Mimeo, Department of Economics, University of Cyprus.

Krueger AB, Lindahl M. 2001. Education for growth: why and for whom? *Journal of Economic Literature* **39**: 1101–1136.

Kyriacou GA. 1991. Level and growth effects of human capital. C.V. Starr Center Economic Research Reports, 91-26, New York University.

Lau LJ, Jamison DT, Louat FF. 1991. Education and productivity in developing countries: an aggregate production function approach. Working Paper No. 612, World Bank.

Li Q, Wang S. 1998. A simple consistent bootstrap test for a parametric regression functional form. *Journal of Econometrics* **87**: 145–165.

Li Q, Huang C, Li D, Fu T. 2002. Semiparametric smooth coefficient models. *Journal of Business and Economic Statistics* **20**: 412–422.

Lucas RE. 1988. On the mechanics of economic development. *Journal of Monetary Economics* **22**: 3–42.

Mankiw N, Romer D, Weil D. 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* **108**: 407–437.

Masanjala WH, Papageorgiou C. 2002. Parameter heterogeneity and nonlinearities in the aggregate production function: investigating the Solow growth model with CES technology. *Journal of Applied Econometrics* (forthcoming).

Nehru V, Swanson E, Dubey A. 1995. A new database on human capital stock in developing and industrial countries: sources, methodology and results. *Journal of Development Economics* **46**: 379–401.

Pissarides CA. 1997. Learning by trading and the returns to human capital in developing countries. *The World Bank Economic Review* **11**: 17–32.

Pritchett L. 1996. Population growth, factor accumulation and productivity. Working Paper No. 1567, World Bank.

Pritchett L. 2001. Where has all the education gone? *World Bank Economic Review* **15**: 367–391.

Psacharopoulos G. 1994. Returns to investment in education: a global update. *World Development* **22**: 1325–1343.
Racine J, Li Q. 2004. Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* **119**: 99–130.
Rodrik D. 2000. Institutions for high-quality growth: what they are and how to acquire them? NBER Working Paper No. 7540.
Sachs J, Warner A. 1995. Economic reform and the process of global integration. *Brookings Papers on Economic Activity* **1**: 1–118.
Stone CJ. 1977. Consistent nonparametric regression. *Annals of Statistics* **5**: 595–620.
Temple J. 1999. The new growth evidence. *Journal of Economic Literature* **37**: 112–156.
Topel R. 1999. Labor markets and economic growth. In *Handbook of Labor Economics*, Vol. 3C, Ashenfelter O, Card D (eds). North Holland: Amsterdam.
Transparency International. 2001. Corruption Perceptions Index: Internet Center for Corruption Research. Available at http://www.gwdg.de/∼uwvw/.
United Nations. *National Accounts Statistics: Main Aggregates and Detailed Tables*. United Nations Publications: New York (various issues).
World Bank. 1989. *World Development Report 1989*. Oxford University Press for the World Bank: New York.
Zhang W, Lee S-Y, Song X. 2002. Local polynomial fitting in semivarying coefficient model. *Journal of Multivariate Analysis* **82**: 166–188.
Zheng JX. 1996. A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* **75**: 263–289.

