Capital–Skill Complementarity?
Evidence from a Panel of Countries*

John Duffy
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260
E–mail: jduffy@pitt.edu

Chris Papageorgiou
Department of Economics
Louisiana State University
Baton Rouge, LA 70803
E–mail: cpapa@lsu.edu

Fidel Perez–Sebastian
Dpto. F. del Análisis Económico
Universidad de Alicante
03690 Alicante, Spain
E–mail: fidel@merlin.fae.ua.es

January 2003

Abstract

Since Griliches (1969), researchers have been intrigued by the idea that physical capital and skilled labor are relatively more complementary than physical capital and unskilled labor. In this paper we consider the cross–country evidence for capital–skill complementarity using a time–series, cross–section panel of 73 developed and less developed countries over a 25 year period. We focus on three empirical issues. First, what is the best specification of the aggregate production technology to address the capital–skill complementarity hypothesis. Second, how should we measure skilled labor? Finally, is there any cross–country evidence in support of the capital–skill complementarity hypothesis? Our main finding is that we find some support for the capital–skill complementarity hypothesis in our macro panel dataset.

JEL Classification Numbers: O40, O47.

*We are grateful to an anonymous referee for very useful suggestions. We also thank Francesco Caselli, Lola Collado, David DeJong, William Green, Daniel Hamermesh, Carter Hill, Jose–Victor Ríos–Rull for their helpful comments and suggestions and Jong–Wha Lee for making available to us his dataset on schooling duration. Financial support from Instituto Valenciano de Investigaciones Económicas and Spanish Ministry of Science and Technology, BEC2001-0980, is gratefully acknowledged.

1 Introduction

Over 30 years ago, Griliches (1969) provided evidence from U.S. manufacturing data suggesting that capital and skilled labor are relatively more complementary as inputs than are capital and unskilled labor. Griliches referred to this finding as the "capital–skill complementarity" hypothesis. Griliches' hypothesis has received renewed attention lately, as the U.S. and other developed nations have invested heavily in "skill–biased" information technology and this development appears to have coincided with a rise in the wages earned by skilled workers relative to the wages of unskilled workers. Indeed, belief in the existence of capital–skill complementarity is so strong that some researchers have suggested modifying the standard neoclassical production technology to account for this phenomenon in addressing questions of economic growth, trade and inequality (see, e.g. Stokey (1996), and Krusell et al. (2000)).

Goldin and Katz (1998) have recently reminded us that physical capital and skilled labor have not always been viewed as relative complements. For example, they note that in an earlier era, the transformation from skilled artisan shops to factories involved the substitution of physical capital and/or unskilled labor for highly skilled labor – precisely the opposite of what is hypothesized to be happening today. Goldin and Katz's findings suggest that capital–skill complementarity may only be a transitory phenomena. As countries progress through various stages of development, skilled labor may change from being relatively more substitutable with capital and unskilled labor to being highly complementary to these two inputs. It therefore seems important to consider the evidence for capital–skill complementarity over long periods of time and across countries at different stages of development. The aim of this paper is to conduct such an exercise. In particular, we examine the evidence for capital–skill complementarity using a panel dataset of 73 countries over the period 1965–1990.

Not surprisingly, since Griliches (1969), the capital–skill complementarity hypothesis has attracted the attention of many researchers who have mainly used cross–sectional manufacturing data for a single (typically) developed country to test this hypothesis. Hamermesh (1993) assesses the findings from most of these studies and concludes that there "may be" capital–skill complementarity. However, he cautions that "many of the studies that disaggregate the work force by demographic group exclude capital as a productive input due to the difficulty of generating satisfactory data on capital stocks in the cross sections examined" (Hamermesh (1993) p. 113). For example, in the original Griliches (1969) study, the assumption of perfectly competitive markets allows data on rates of return to proxy for the marginal product of capital and thereby capture variations in the stock of capital. By contrast, in this paper, we make use of the Penn World Tables–Version 5.6 dataset on investment rates across countries to construct physical capital stocks. We examine the capital–skill complementarity hypothesis *directly*, without resorting to assumptions of perfectly competitive markets, by estimating the parameters of various different specifications of an aggregate production function.[1] While the competitive markets assumption may seem reasonable for developed countries, it may be less reasonable for developing countries where factors may be less mobile and markets less complete.

Hamermesh (1993) also notes the difficulties that earlier studies had in using occupational data to differentiate between skilled and unskilled workers. In this paper, we follow the tradition in the macro–growth literature and differentiate labor according to educational attainment levels using the recent Barro and Lee (2001) dataset. In particular, we consider five alternative proxies for skilled labor ranging from workers possessing some primary education to workers possessing some post–secondary education; for each proxy, the fraction of the labor force that does not meet the educational threshold used to define skilled labor is regarded as unskilled labor. We also examine what happens when we augment our labor data with data on returns to schooling (earnings) in an effort to account for disparities in efficiency units across workers *within* the class of workers regarded as skilled or unskilled. Our analysis of several different classifications and measures of skilled and unskilled labor is another novel feature of this study; in prior studies involving skilled and unskilled labor, a single educational threshold has been chosen to divide workers into skilled and unskilled classes without much consideration being given to the empirical relevance of the threshold choice.

International examinations of the capital–skill complementarity hypothesis have been conducted by Fallon and Layard (1975), Berman et al. (1998) and Flug and Hercowitz (2000). Our approach is most closely related to the Fallon and Layard (1975) study; the Berman et al. and Flug and Hercowitz studies do not employ aggregate production functions to test capital–skill complementarity across countries. Fallon and Layard used data pieced together for 9 developed and 13 less developed countries for a single year, 1963, to estimate reduced form equations derived from two–level CES production functions that allowed for there to be differences in the elasticity of substitution

[1]The methodology used in this paper follows Duffy and Papageorgiou (2000) who investigate a general two–factor CES aggregate specification in which output is generated using physical capital and labor or human capital adjusted labor serving as inputs.

between capital and skilled labor and the elasticity of substitution between capital and unskilled labor. At the economy–wide level, they find "mild" (though statistically insignificant) evidence in favor of the capital–skill complementarity hypothesis. In this paper, we also make use of the two–level CES production function specification that Fallon and Layard advocate. However, since we use nonlinear estimation methods that were not feasible at the time of the Fallon and Layard study, we do not need to follow Fallon and Layard further in assuming perfectly competitive markets so that factor price data (reflecting marginal products under perfect competition) can be used to estimate linear reduced form equations. Furthermore, we use data for many more countries, 73, and there is also a time dimension to our panel dataset that was missing from Fallon and Layard's study. Specifically, for each of the 73 countries, we have 6 annual observations, spaced five years apart: 1965,1970,...,1990 (a total of 438 observations). Our analysis thus allows for a clearer and more convincing assessment of whether the capital–skill complementarity hypothesis is common to many countries over some length of time.

Our main finding is that there is indeed some evidence in support of the capital–skill complementarity hypothesis. This hypothesis finds greatest support when skilled labor is defined using a low threshold, e.g. classifying skilled workers as those who have completed a *primary* education. This threshold for skilled labor is considerably lower than the threshold that other researchers have used to define skilled labor, e.g. classifying skilled workers as those who have more than a *secondary* education (e.g. Krusell et al. (2000)). Still, we emphasize that our evidence in favor of the capital–skill complementarity hypothesis is weak. Additional caution in the interpretation of our results comes from the Monte Carlo experiments that we perform in the paper. These experiments suggest that, for small sample sizes comparable to what have available in our macro panel dataset, the nonlinear techniques that use to estimate the two–level CES production function produce relatively imprecise estimates. Therefore, our evidence for capital–skill complementarity may not warrant modifying the specifications of the aggregate production technology to take account of this hypothesis. Alternatively our results might reasonably serve to bolster alternative (and complementary) explanations for rising wage and income inequality, for example, skill–biased technological change or country–specific government policies.

2 Examining The Case for Capital–Skill Complementarity using Aggregate Production Functions

The capital–skill complementarity hypothesis states that physical capital is more complementary to skilled labor than to unskilled labor. More formally, suppose aggregate output, Y, is given by a three–factor production technology $Y = F(K, S, N)$, where K denotes the physical capital stock, S denotes the quantity of skilled labor and N denotes the quantity of unskilled labor. Denote by $\sigma_{i,j}$ the elasticity of substitution between inputs i and j. Then capital–skill complementarity holds if $\sigma_{K,N} > \sigma_{K,S}$.

In order to assess the extent of capital–skill complementarity, we must work with a functional form that is general enough to accommodate different elasticities of substitution. For example, the relatively general CES form for $F(K, S, N)$,

$$Y = A\left[aK^{\rho} + bS^{\rho} + cN^{\rho}\right]^{\frac{1}{\rho}},$$

where $a > 0, b > 0, c > 0$ and $\rho \leq 1$, implies that the elasticity of substitution between any two inputs, $\sigma_{i,j}$ for $i, j \in \{K, S, N\}$, is constant and equal to $\frac{1}{1-\rho}$. To allow for different elasticities of substitution between any two inputs we make use of Sato's (1967) two–level CES production function. The two most interesting versions of this two–level CES form for purposes of testing the capital–skill complementarity hypothesis are

$$Y = A\left[a[bK^{\theta} + (1-b)S^{\theta}]^{\rho/\theta} + (1-a)N^{\rho}\right]^{1/\rho}, \tag{1}$$

$$Y = A\left[a[bK^{\theta} + (1-b)N^{\theta}]^{\rho/\theta} + (1-a)S^{\rho}\right]^{1/\rho}, \tag{2}$$

where A is a positive (factor neutral) efficiency parameter, a, b are distribution parameters and $\theta, \rho \leq 1$ are the intra– and inter–class elasticity of substitution parameters, respectively (θ, $\rho = 1$ imply perfect substitutability, θ, $\rho = 0$ imply the Cobb–Douglas specification, and θ, $\rho = -\infty$ imply perfect complementarity). Even though the two specifications are similar, they differ in one important way. In (1), the elasticity of substitution between K and N, and N and S are the same, while in equation (2) the elasticity of substitution between K and S, and N and S are the same. Thus, the capital skill complementarity hypothesis $\sigma_{K,N} > \sigma_{K,S}$ is readily tested using either specification. In particular, as we will demonstrate later, capital–skill complementarity holds in specification (1)[2] iff $\rho > \theta$ [$\rho < \theta$].

Though further disaggregation is possible, e.g. through the use of a translog specification (see, e.g. Bergström and Panas (1992) and Ruiz–Arranz (2002)), we focus on the two–level CES specifications as they are the ones that have been used in the recent literature examining the consequences of the capital–skill complementarity hypothesis. For example, Fallon and Layard (1975) and Caselli and Coleman (2002a,b) both prefer to work with specification (1). Krusell et al. (2000) consider an expanded version of specification (1)

$$Y = AK_s^{\alpha}\left[a[bK_e^{\theta} + (1-b)S^{\theta}]^{\rho/\theta} + (1-a)N^{\rho}\right]^{\frac{1-\alpha}{\rho}},$$

where K_s represents the stock of capital structures, and K_e represents the stock of capital equipment. While we would like to estimate such a specification, we lack the requisite data on capital structures and capital equipment for all of the countries in our sample.[2]

Stokey (1996), on the other hand, has proposed a more restrictive version of specification (2)

$$Y = A[bK^{\theta} + (1-b)N^{\theta}]^{\gamma/\theta}\tilde{S}^{(1-\gamma)}. \tag{3}$$

Here $\tilde{S} = S + qN$ represents "mental effort", $q < 1$ is the relative efficiency of unskilled labor in contributing to mental effort, and $1-\gamma$ is the share of output that accrues to $\tilde{S}$. Equation (3) is clearly a restricted form of (2) as it requires finding that estimates of ρ are not significantly different from zero. Conditional on this finding capital–skill complementarity holds if $0 < \theta \leq 1$.[3]

Goldin and Katz (1998) start off with the two–level CES specification (1) but further specialize it to the case where 1) $\theta \to -\infty$ and 2) $\rho \to 0$. This is even more restrictive than Stokey (1996), since it implies, as in Stokey, that final output Y has the Cobb–Douglas form but it further requires that the K–S aggregate, which Goldin and Katz refer to as K^*, have the Leontief form

$$Y = A\left[(\min[bK, (1-b)S])^{\gamma} N^{1-\gamma}\right].$$

[2]Krusell et al. (2000) only consider the U.S. economy, for which such data are available. They use a two–level CES function, calibrated so that there is capital–skill complementarity, and find that variations in factor inputs can account for most of the variation in the skill premium in the U.S.. Ruiz–Arranz (2002) follows up on Krusell et al. (2000) and examines the effect of capital–skill complementarity on U.S. skill premium between 1965–1999. By contrast with Krusell et al., Ruiz–Arranz's estimation is done using a translog production function that allows her to disentangle the effects of capital–skill complementarity and of skill–biased technological change on the U.S. skill premium. Ruiz–Arranz finds that capital–skill complementarity can account for at most 40 percent of the rise in the skill premium, with skill–biased technological change accounting for most of the variation in the skill premium.

[3]Following Stokey's formulation, the restricted version of the two–level CES specification (1) is

$$Y = A[bK^{\theta} + (1-b)S^{\theta}]^{\gamma/\theta}N^{1-\gamma},$$

and capital–skill complementarity holds if $\theta < 0$.

In this case, since $\sigma_{K,S} = 0 < 1$ and $\sigma_{K^*,N} = 1$, the authors are making the empirically testable assumption that $\sigma_{K,S} < \sigma_{K^*,N}$. Their aim is to show that if technology changes, represented by a change in A, then it need not be the case that the relative demand for skilled labor increases. As A increases, less is needed of both the K^* aggregate and N to produce the same level of output.

2.1 Elasticity of Substitution Measures

For general production technologies with more than two inputs there is no single definition for the elasticity of substitution between pairs of inputs. Perhaps the most commonly used definition is the Allen–Uzawa *partial* elasticity of substitution that measures the percentage change in the ratio of two inputs in response to a change in the ratio of the two input prices, holding all other prices (but not all other inputs) and output quantity constant. This is the measure used, e.g. by Griliches (1969). Another elasticity of substitution definition is the Hicks–Allen *direct partial* elasticity of substitution that measures the percentage change in the ratio of two inputs in response to a change in the ratio of the two input prices, holding all other prices, inputs and output quantity constant.[4] In what follows, we show that in the two–level CES specification (1)[2], the capital–skill complementarity hypothesis ($\sigma_{K,N} > \sigma_{K,S}$) holds iff $\rho > \theta$ [$\rho < \theta$] regardless of which elasticity measure we use, the Allen partial elasticity of substitution or the direct partial elasticity of substitution.[5]

2.1.1 The Allen Partial Elasticity of Substitution

Suppose there are n inputs used in production. Let $[X_1, ..., X_m]$ form a partition of these inputs into $m \leq n$ distinct subsets or input *classes*. For example, in specification (1) we have $m = 2$ with X_1 consisting of K and S while X_2 consists of N alone. As shown in Sato (1967 pp. 202–204, 216–217) the general two–level CES production function implies that the Allen partial elasticity of

[4] More formally, assume that output Y is produced using n inputs, $x = \{x_1, ..., x_n\}$, according to some general production technology $Y = F(x)$. Letting $F_i \equiv \partial F/\partial x_i$ and $F_{ij} \equiv \partial^2 F/\partial x_i \partial x_j$, the Allen partial elasticity of substitution is defined by $\sigma^a_{ij} = \frac{\mathcal{H}_{ij} \sum_{k=1}^n x_k F_k}{x_i x_j |\mathcal{H}|}$, where

$$\mathcal{H} \equiv \begin{bmatrix} 0 & F_1 & \dots & F_n \\ F_1 & F_{11} & \dots & F_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ F_n & F_{n1} & \dots & F_{nn} \end{bmatrix}$$

and $\mathcal{H}_{ij}$ is the cofactor of element F_{ij} in $\mathcal{H}$. The direct partial elasticity of substitution between the ith and the jth element in $\{x\}$ is given by $\sigma^d_{ij} = -\frac{\partial \ln(x_i/x_j)}{\partial \ln(F_i/F_j)}$, where y and all x_k are held constant.

[5] We thank an anonymous referee for pointing out the alternative definitions of the elasticity of substitution and their potential effect on capital–skill complementarity.

substitution between the ith and jth inputs is given by

$$\sigma_{ij}^{a} = \begin{cases} \sigma, & \text{if } i \in X_r, j \in X_s, r \neq s \\ \sigma_s + \frac{1}{\zeta^s}(\sigma_s - \sigma), & \text{if } i, j \in X_s, i \neq j \end{cases}$$

where σ is the inter–class elasticity of substitution, σ_s is the intra–class elasticity of substitution within class s, and ζ^s is the relative share of the factors in class s in total expenditure. In the case of the two–level CES specification (1) we have

$$\begin{aligned} \sigma_{K,N}^{a} &= \sigma_{N,S}^{a} = \frac{1}{1-\rho} \\ \sigma_{K,S}^{a} &= \frac{1}{1-\rho} + \frac{1}{\zeta^s}\left(\frac{1}{1-\theta} - \frac{1}{1-\rho}\right), \end{aligned}$$

where ζ^s is the relative share of K and S in total expenditure.[6] Imposing the capital–skill complementarity condition gives

$$\begin{aligned} & \sigma_{K,N}^{a} > \sigma_{K,S}^{a} \\ \Leftrightarrow\quad & \frac{1}{1-\rho} + \frac{1}{\zeta^s}\left(\frac{1}{1-\theta} - \frac{1}{1-\rho}\right) < \frac{1}{1-\rho} \\ \Rightarrow\quad & \frac{1}{\zeta^s}\left(\frac{1}{1-\theta} - \frac{1}{1-\rho}\right) < 0. \end{aligned}$$

Given that $\zeta^s \in (0,1)$ it follows that $\rho > \theta$.[7] Following the same logic, one can show that for specification (2), $\sigma_{K,N} > \sigma_{K,S}$ holds iff $\rho < \theta$.

2.1.2 The Direct Partial Elasticity of Substitution

Following Sato (1967 pp. 202–204, 216–217), for the general two–level CES production function, the direct partial elasticity of substitution between the ith and jth inputs is given by

$$\sigma_{ij}^{d} = \begin{cases} \dfrac{\frac{1}{\zeta_i^r} + \frac{1}{\zeta_j^s}}{\frac{1}{\sigma_r}\left(\frac{1}{\zeta_i^r} - \frac{1}{\zeta^r}\right) + \frac{1}{\sigma_s}\left(\frac{1}{\zeta_j^s} - \frac{1}{\zeta^s}\right) + \frac{1}{\sigma}\left(\frac{1}{\zeta^r} + \frac{1}{\zeta^s}\right)}, & \text{if } i \in X_r, j \in X_s, r \neq s \\ \sigma_s, & \text{if } i, j \in X_s, i \neq j \end{cases} \tag{4}$$

[6]Notice that the Allen partial elasticity of substitution between K and S is *not* constant.

[7]Fallon and Layard (1975) also use two–level CES productions functions. They relate the capital–skill complementarity condition for the two–level specification (1) (i.e. $\sigma_{K,N} > \sigma_{K,S} \Rightarrow \rho > \theta$) to Hicks partial elasticity of *complementarity* concept for *general* production function specifications, which is essentially the dual of the Allen partial elasticity of substitution. The partial elasticity of complementarity measures the percentage change in the ratio of two input prices in response to a change in the ratio of the two input quantities holding all other inputs and the price of output constant. It is given by: $c_{ij} = \frac{FF_{ij}}{F_iF_j}$. Using this elasticity measure, capital-skill complementarity requires $c_{S,K} > c_{N,K}$. Fallon and Layard (1975) show that for the two–level CES specification (1), $c_{S,K} = 1 - \rho + \frac{1}{\zeta^s}(\rho - \theta)$ where, as above, ζ^s is the relative share of K and S in total expenditure and $c_{N,K} = 1 - \rho$. Hence, capital–skill complementarity, $c_{S,K} > c_{N,K}$, requires again that $\rho > \theta$.

where σ_s and σ_r are the *intra*–class elasticities of substitutions between pairs of inputs within classes s and r, respectively, while σ is the *inter*–class elasticity of substitution, $\zeta^s(\zeta^r)$ is the relative expenditure share of the sth (rth) class of factors and $\zeta_j^s(\zeta_i^r)$ is the relative share of the jth (ith) element of the sth (rth) class in total expenditure.

As noted above, the two–level CES specification (1) consists of only two classes – one (representing the first level) that aggregates two inputs (K and S) using CES, and another that employs only one input (N) using Cobb–Douglas – nested in another CES (representing the second level). For specification (1), we have $\sigma_s = \frac{1}{1-\theta}, \sigma = \frac{1}{1-\rho}$. Since the second class in our production function is Cobb–Douglas it follows that $\sigma_r = 1$ and $\frac{1}{\zeta_i^r} = \frac{1}{\zeta^r}$. According to equation (4),

$$\begin{aligned}\sigma_{K,S}^d &= \frac{1}{1-\theta}\\ \sigma_{K,N}^d &= \sigma_{N,S}^d = \frac{\frac{1}{\zeta^r} + \frac{1}{\zeta_j^s}}{(1-\theta)\left(\frac{1}{\zeta_j^s} - \frac{1}{\zeta^s}\right) + (1-\rho)\left(\frac{1}{\zeta^r} + \frac{1}{\zeta^s}\right)}.\end{aligned}$$

Imposing the capital–skill complementarity condition gives[8]

$$\sigma_{K,N}^d > \sigma_{K,S}^d$$

$$\begin{aligned}&\Leftrightarrow \quad \frac{1}{1-\theta} < \frac{\frac{1}{\zeta^r} + \frac{1}{\zeta_j^s}}{(1-\theta)\left(\frac{1}{\zeta_j^s} - \frac{1}{\zeta^s}\right) + (1-\rho)\left(\frac{1}{\zeta^r} + \frac{1}{\zeta^s}\right)}\\ &\Rightarrow \quad (1-\theta)\left(\frac{1}{\zeta^r} + \frac{1}{\zeta^s}\right) > (1-\rho)\left(\frac{1}{\zeta^r} + \frac{1}{\zeta^s}\right)\\ &\Rightarrow \quad \rho > \theta.\end{aligned}$$

Again following the same logic, one would conclude that for specification (2), $\sigma_{K,N} > \sigma_{K,S}$ holds iff $\rho < \theta$.

In summary, regardless of the way in which the elasticity of substitution is defined, capital–skill complementarity in specification (1) [2] holds iff $\rho > \theta$ [$\rho < \theta$].

3 Estimation Procedures and Specifications

While there is some supporting evidence for the capital–skill complementarity hypothesis using alternative datasets and methodologies as noted in the introduction, the hypothesis has not been

[8]Once again, notice that the direct partial elasticities of substitution between K and N and between N and S are *not* constant.

tested 1) using aggregate production function specifications *directly* or 2) using a cross–section, time–series panel dataset.[9] The latter point is particularly relevant in growth models that use the aggregate production functions motivated by the supposed existence of capital–skill complementarities. In addition, as our literature review suggested, there is no consensus yet on the appropriate functional form to use to capture capital–skill complementarity. Our estimation exercise, to which we now turn, sheds some light on this question as well.

The two–level CES specifications we consider are highly nonlinear and therefore, nonlinear estimation methods (in particular, nonlinear least squares (NLLS) and generalized method of moments (GMM) estimators) will be used to obtain estimates of ρ and θ. These computationally intensive methods were not feasible when Fallon and Layard (1975) first proposed estimation of production function specifications, and consequently, they had to resort to estimation of restrictive linear specifications as noted in the introduction.

3.1 The Two–Level CES Specifications

The two–level CES production function equations that will be empirically tested are

$$Y_{it} = A_{i0}\left[a[bK_{it}^{\theta} + (1-b)S_{it}^{\theta}]^{\rho/\theta} + (1-a)N_{it}^{\rho}\right]^{1/\rho} e^{\lambda t + \varepsilon_{it}}, \tag{5}$$

$$Y_{it} = A_{i0}\left[a[bK_{it}^{\theta} + (1-b)N_{it}^{\theta}]^{\rho/\theta} + (1-a)S_{it}^{\rho}\right]^{1/\rho} e^{\lambda t + \varepsilon_{it}}, \tag{6}$$

where i denotes the country, t denotes the year and ε is the error term. We assume exogenous, Hicks–neutral technological growth. In particular, we assume A is growing at the rate λ, with A_{i0} representing the initial ($t = 0$) value of A for country i.[10] Notice that model specification (5) corresponds to the first version of the two–level CES form, equation (1), and model specification (6) corresponds to the second version of the two–level CES form, equation (2). While it is possible to linearize equations (5–6), the resulting equations are complicated and impossible to estimate.[11] The only remaining viable option is nonlinear estimation and that is how we proceed.

[9] Flug and Hercowitz (2000) who investigate the related idea of an equipment–skill complementarity hypothesis do use international panel data from 35 countries. However, they do not estimate production functions directly as we do here. Instead, they use a linear regression model of wage and unemployment ratios of skilled to unskilled workers. Their results suggest that investment in equipment raises the relative demand for skilled workers.

[10] That is, $A_{it} = A_{i0}e^{\lambda t}$. In an interesting paper, Caselli and Coleman (2002a) use a two–level CES specification in which they allow the efficiency parameters for the three different factors, unskilled labor, skilled labor and capital to differ from one another.

[11] Using a second order Taylor series expansion it is possible to obtain a linear approximation of the two–level CES specification. Unlike the linearized version of Stokey's formulation, discussed below (in footnote 13), the linearized approximation of the two–level CES specification (linearized around $\rho, \theta = 0$) contains a large number of linear parts with multiple coefficients that cannot be identified using standard linear estimation techniques.

In using panel data for our estimation exercise, we must confront two potential econometric problems. First, there is the problem of unmodeled, country specific *fixed–effects*, due for example, to differences in technology, culture or geography (see, e.g. Islam (1995)). Assuming these factors are time invariant, we can resolve the fixed effects problem by supposing that the error term, $\varepsilon_{it} = \eta_i + \epsilon_{it}$, where η_i represents the country specific fixed factors in country i. Under this assumption, log differencing (5) and (6) yields

$$\log\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = \lambda + \frac{1}{\rho}\log\frac{\left[a[bK_{it}^{\theta} + (1-b)S_{it}^{\theta}]^{\rho/\theta} + (1-a)N_{it}^{\rho}\right]}{\left[a[bK_{i,t-1}^{\theta} + (1-b)S_{i,t-1}^{\theta}]^{\rho/\theta} + (1-a)N_{i,t-1}^{\rho}\right]} + \epsilon_{it} - \epsilon_{i,t-1}, \quad (7)$$

$$\log\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = \lambda + \frac{1}{\rho}\log\frac{\left[a[bK_{it}^{\theta} + (1-b)N_{it}^{\theta}]^{\rho/\theta} + (1-a)S_{it}^{\rho}\right]}{\left[a[bK_{i,t-1}^{\theta} + (1-b)N_{i,t-1}^{\theta}]^{\rho/\theta} + (1-a)S_{i,t-1}^{\rho}\right]} + \epsilon_{it} - \epsilon_{i,t-1}. \quad (8)$$

A second problem concerns the possible endogeneity of the input variables in our regression specifications, as emphasized by Caselli et al. (1996). We resolve this second problem by using a GMM, instrumental variables procedure to estimate the log–differenced model, where we use suitable lagged values of the input and output variables as instruments.

3.2 CES–nested–in–Cobb–Douglas Specification

An alternative to the two–level CES specifications is the more restricted version of these specifications proposed by Stokey (1996) as given by equation (3). Our estimated version of Stokey's production function specification is of the following form:

$$Y_{it} = A_{i0}[bK_{it}^{\theta} + (1-b)N_{it}^{\theta}]^{\gamma/\theta} S_{it}^{1-\gamma} e^{\lambda t + \varepsilon_{it}}. \quad (9)$$

In (9), capital and unskilled workers are combined into an aggregate by a CES specification. The resulting aggregate measure is then combined with skilled labor using a Cobb–Douglas technology. Notice that our specification (9) is really a special case of (3) in that we assume that $q = 0$; this assumption implies that mental effort in the production process is exerted only by skilled workers.[12] The capital–skill complementarity would hold in this case if the elasticity of substitution between capital and unskilled workers is greater than unity, that is, $0 < \theta \leq 1$. Similarly, the restricted version of specification (1) that we will estimate is given by

$$Y_{it} = A_{i0}[bK_{it}^{\theta} + (1-b)S_{it}^{\theta}]^{\gamma/\theta} N_{it}^{1-\gamma} e^{\lambda t + \varepsilon_{it}}, \quad (10)$$

[12] There exists no empirical evidence on q (the contribution of unskilled labor to mental effort). Stokey (1996) simply assumes that $q = 0.25$ in order to keep the skill premium within a reasonable range in her calibration exercises.

where the sufficient condition for capital–skill complementarity is reversed, $\theta < 0$. We will refer to specifications (9–10) as the "CES–nested–in–Cobb–Douglas" specifications, and we will estimate them using nonlinear least squares.

As in the case of the general, two–level CES specifications, we also consider a log–difference version of the CES–nested–in–Cobb–Douglass specification that gets rid of country–specific fixed effects. Log–differencing (10) and (9) (note the change in order) we obtain the following two expressions:

$$\log\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = \lambda + \frac{\gamma}{\theta}\log\frac{[bK_{it}^{\theta} + (1-b)S_{it}^{\theta}]}{[bK_{i,t-1}^{\theta} + (1-b)S_{i,t-1}^{\theta}]} + (1-\gamma)\log\left(\frac{N_{it}}{N_{i,t-1}}\right) + \epsilon_{it} - \epsilon_{i,t-1}, \quad (11)$$

$$\log\left(\frac{Y_{it}}{Y_{i,t-1}}\right) = \lambda + \frac{\gamma}{\theta}\log\frac{[bK_{it}^{\rho} + (1-b)N_{it}^{\theta}]}{[bK_{i,t-1}^{\theta} + (1-b)N_{i,t-1}^{\theta}]} + (1-\gamma)\log\left(\frac{S_{it}}{S_{i,t-1}}\right) + \epsilon_{it} - \epsilon_{i,t-1}. \quad (12)$$

We will estimate (11–12) using nonlinear least squares and using a GMM, instrumental variables procedure where lagged values of input and output variables are used as instruments.[13]

4 The Data

Our estimation requires data for real GDP (Y), the stock of physical capital (K), unskilled labor (N), and skilled labor (S). We obtain data for Y from the Penn World Tables v. 5.6 (PWT–5.6), and construct data for K using investment shares data from the PWT–5.6 and the perpetual inventory approach. Data for both Y and K are in constant U.S. dollars (1985 international prices). Since the data we use to construct the skilled labor proxies are only available every five years, our dataset consists of a number of annual observations (6) for each country, spaced five years apart. We constructed five alternative proxies for skilled (unskilled) labor since it was not clear to us how skilled labor should be defined. Our five proxies for skilled labor are: 1) workers who have attained

[13]We note that it is possible to obtain a linearized version of the restricted CES–nested–in–Cobb Douglas specification. Divide the left and right hand sides of (10) by N_{it}, and the left and right hand sides of (9) by S_{it}. Log–linearizing the resulting equations around $\theta = 0$ gives respectively

$$\log y_{it} = \log A_{i0} + \lambda t + \gamma b \log k_{it} + \gamma(1-b)\log s_{it} + 1/2\gamma b(1-b)\theta\left(\log\frac{k_{it}}{s_{it}}\right)^2 + \varepsilon_{it},$$

where $y = \frac{Y}{N}$, $k = \frac{K}{N}$, $s = \frac{S}{N}$ and

$$\log y_{it} = \log A_{i0} + \lambda t + \gamma b \log k_{it} + \gamma(1-b)\log n_{it} + 1/2\gamma b(1-b)\theta\left(\log\frac{k_{it}}{n_{it}}\right)^2 + \varepsilon_{it},$$

where $y = \frac{Y}{S}$, $k = \frac{K}{S}$, $n = \frac{N}{S}$. We obtained estimates from these linear specifications using OLS with time and fixed effects and instrumental variables but found that they did not change the main conclusions we obtained from the more general nonlinear specifications. We therefore chose to omit these findings from the paper.

some post–secondary (college) education (labeled $S1$), 2) workers who have completed secondary education ($S2$), 3) workers who have attained some secondary education ($S3$), 4) workers who have completed primary education, ($S4$), and 5) workers who have attained some primary education ($S5$).[14] Each skilled labor proxy was constructed by multiplying achievement rates for a particular cut–off criterion (using data from Barro and Lee (2001)) by the size of the labor force in each country at each date in our sample. The remainder of the labor force (those not classified according to the definition of skilled labor ($S1$–$S5$)) was regarded as unskilled labor, and was designated by $N1$, $N2$, $N3$, $N4$ or $N5$, corresponding to the definition of skilled labor. Our balanced panel dataset consists of 73 countries; for each country there are six annual observations of all input and output variables spaced five years apart starting in 1965 and ending in 1990 (438 observations). We choose to work with a large panel of countries, rather than estimating production functions for individual countries as we have only six observations per country and the CES specifications involve as many as six parameters.

Since workers who have attained some college education may contribute more efficiency units than workers who have only attained some secondary education, the proxies we used for skilled (unskilled) labor could suffer from aggregation problems, for example, when skilled labor is defined as those who have attained secondary education ($S3$). In an effort to address this problem, we follow Caselli and Coleman (2002a) and employ additional data on returns to schooling to weight individuals *within* our two divisions of the labor force into skilled or unskilled labor. We will refer to this dataset as the "weighted" labor data to differentiate it from the data where returns to schooling data are not used in the construction of proxies for skilled and unskilled labor (the "unweighted" labor data). While adjusting the skilled/unskilled labor proxies to account for returns to schooling may seem quite reasonable, it comes at the cost of drastically reducing our sample size from 73 to 49 countries (from 438 to 294 observations) due to the lack of data on returns to schooling for 24 countries. We will return to this issue later in the paper. Because of this data constraint, we report results for both the larger, unweighted labor dataset and the smaller weighted labor dataset.

Appendix 1 provides further details concerning the sources and construction of the data used in this paper as well as a table reporting the mean values of Y, K, $S4$ and $N4$ for each country in the sample.

[14]In an earlier draft of this paper we considered a definition of skilled labor as comprising those who had *completed* post–secondary (college) education. However, in many less developed countries, the fraction of workers meeting this definition was close to zero, and so we dropped this definition from our analysis.

5 Results

Our results are organized as follows. We first address the question concerning which specification, (1) or (2), is preferred. We then report estimation results for the preferred specification using the various estimation techniques; without and with fixed effects removed (with FE) and using instrumental variable (IV) estimators. We also consider the robustness of our specification and estimation results using additional data on wage rates to augment our measures of skilled labor. Finally, we report the results of a Monte Carlo exercise that allows us to assess the reliability of the parameter estimates we report in the paper. We proceed by first reporting our estimation results obtained from using the unweighted–labor data and then commenting on the respective results obtained from using the weighted–labor data (the latter results are qualitatively similar to those obtained using the unweighted data and hence are presented in Appendix 4).

5.1 Specification Search

The two main competing specifications for testing the capital–skill complementarity hypothesis are given by our equations (1) and (2). Within each of these two specifications, we considered the two–level nonlinear model and the CES–nested–in–CD model without or with fixed effects removed. For regression models based on specification (2) we frequently obtained parameter estimates that had the wrong signs, had very large standard errors or were empirically implausible in magnitude, e.g. estimates for ρ in excess of one.[15] By contrast, using specification (1) our estimated parameters generally have the right signs and are almost always empirically plausible in magnitude though not necessarily statistically significant. For this reason we prefer specification (1) to specification (2); the rest of the paper reports and analyzes results from estimation models based on specification (1) alone. A more extensive justification for our choice of specification (1) over (2) is given in Appendix 2, which reports some regression results using specification (2). Within our preferred specification (1), the question that remains is the appropriate estimation model, that is, whether to use the two–level CES or the CES–nested–in–CD. We now turn our attention to addressing this question.

[15] Fallon and Layard (1975 Appendix A) and Krusell et al. (2000) report similar difficulties with estimates from versions of specification (2).

5.2 Coefficient Estimates

Table 1 presents coefficient estimates obtained from nonlinear regressions using the unweighted–labor data in various versions of specification (1). All of the NLLS estimation results reported in Table 1 (and subsequent tables) were obtained using economically plausible initial parameters. A grid search on the initial parameter values was also conducted to assess the robustness of the results.

Under the column in Table 1 labeled "NLLS," we report nonlinear least squares (NLLS) parameter estimates for specification (5) (the two–level model with no correction for fixed effects) for each of the five ways of classifying skilled labor. Under the column "NLLS with FE," we report NLLS estimates for the log difference specification (7) (the two–level model with fixed effects (FE) removed) again for all five ways of classifying skilled labor. Finally, under the column "GMM–IV with FE" we report estimates from a GMM–IV procedure applied to the log–difference specification (7). The GMM–IV estimator was chosen to deal with a possible endogeneity problem arising from the fact that the error term in the log–difference specification (7) is likely to be contemporaneously correlated with the input variables, K_{it}, S_{it} and N_{it}. More generally, the perpetual inventory approach used to construct capital stock values (see the appendix for details) implies that K_{it} will always depend on $\epsilon_{i,t-1}$, which is one of the two components of the log–difference specification error term at date t. Notice that $\epsilon_{i,t-1}$ is also a component of this disturbance at time $t-1$; which implies that the error term in the log–difference specification displays first–order serial correlation. To address these possible endogeneity problems, we also use a GMM–IV estimator to estimate the first–difference (fixed effects removed) specifications. Following the framework outlined in Arellano and Bond (1991), our first–difference GMM–IV estimator uses as instruments input–output variables lagged two periods or more, and corrects for the first–order autocorrelated disturbances.[16] This methodology was initially imported into the growth literature by Caselli et al. (1996) and has subsequently become an important benchmark estimation method.[17]

[16]In particular, our GMM estimation of (7) and (11) (results from the latter are presented later in Table A3) uses $\log Y_{i,t-2}$, $\log Y_{i,t-3}$, $\log K_{i,t-2}$, $\log K_{i,t-3}$, $\log S_{i,t-2}$, $\log S_{i,t-3}$ and $\log N_{i,t-2}$, $\log N_{i,t-3}$ as instruments. Notice that the number of observations we have available for our GMM regressions is greatly reduced by the use of three lags of all variables. Sensitivity analyses were performed using alternative sets of instruments, e.g. the smaller set $\log K_{i,t-2}$, $\log K_{i,t-3}$, $\log S_{i,t-2}$, $\log S_{i,t-3}$ and $\log N_{i,t-2}$ $\log N_{i,t-3}$, and not correcting for first-order autocorrelation. We do not report these results as they are similar to those reported in Tables 1 and A3.

[17]Blundell and Bond (1998, 2000) suggest an alternative approach that involves GMM estimation of a *system* of production functions in both levels and first differences using lagged first differences of all variables dated $t-2$ and earlier as instruments in the levels equation and lagged levels dated $t-3$ and earlier as instruments in the first

Table 1: Two–Level CES Nonlinear Estimation

Skilled Labor	Parameter	NLLS	NLLS with FE	GMM–IV with FE
Attained	ρ	0.54638***	0.23861***	0.78668***
		(0.06839)	(0.07568)	(0.29509)
College	θ	0.20459	0.52216	0.45567
		(0.15939)	(1.01680)	(2.4695)
	$\rho - \theta$	0.34179*	-0.28355	0.33101
		(0.19206)	(1.0419)	(2.3159)
	LR Test $\rho = \theta$	2.3612	0.36876	—
	$p > \chi^2(df = 1)$	0.12439	0.54368	—
Completed	ρ	0.54344***	0.37839***	1.2575***
		(0.07226)	(0.07248)	(0.46188)
Secondary	θ	0.43718**	0.50824	0.15426
		(0.20005)	(0.29852)	(1.7342)
	$\rho - \theta$	0.10626	-0.12985	1.1033
		(0.20522)	(0.30483)	(1.5256)
	LR Test $\rho = \theta$	0.20921	0.19729	—
	$p > \chi^2(df = 1)$	0.64739	0.65692	—
Attained	ρ	0.59841***	0.50364***	0.76559
		(0.08194)	(0.07467)	(0.95709)
Secondary	θ	0.45194***	-0.07194	0.56737
		(0.16511)	(0.20993)	(1.9914)
	$\rho - \theta$	0.14647	0.57559***	0.19822
		(0.19111)	(0.20739)	(1.9886)
	LR Test $\rho = \theta$	0.50606	8.8075	—
	$p > \chi^2(df = 1)$	0.47685	0.00210	—
Completed	ρ	0.64170***	0.66110***	0.83540
		(0.09228)	(0.09904)	(0.52395)
Primary	θ	0.51848***	-0.04502	0.32611
		(0.18319)	(0.17359)	(4.7743)
	$\rho - \theta$	0.12321	0.70612***	0.50929
		(0.23352)	(0.19696)	(4.7923)
	LR Test $\rho = \theta$	9.9200	13.636	—
	$p > \chi^2(df = 1)$	0.00163	0.00022	—
Attained	ρ	0.90755***	0.80162***	0.89721
		(0.15607)	(0.14453)	(1.2219)
Primary	θ	0.28622**	0.13594	-0.03832
		(0.13775)	(0.11820)	(3.0498)
	$\rho - \theta$	0.62134***	0.66567***	0.93553
		(0.22656)	(0.17395)	(4.0566)
	LR Test $\rho = \theta$	4.4323	15.952	—
	$p > \chi^2(df = 1)$	0.03761	0.00007	—
No. Obs.		438	365	219

Notes: Standard errors, in parentheses, were recovered using standard approximation methods. White's heteroskedasticity correction was used. ***, **, * Significantly different from 0 at the 1% , 5%, and 10% levels, respectively.

Standard NLLS estimation of specification (5) without the removal of fixed effects or use of instruments – see the "NLLS" column of Table 1 – yields estimates for ρ and θ that are positive and, with one exception, significantly different from zero. Recall that for this two–level specification, capital–skill complementarity is said to obtain if $\rho > \theta$. The NLLS estimates suggest capital–skill complementarity for our two extreme definitions of skilled labor; the difference $\rho - \theta$ is found to be significantly positive when skilled labor is defined as those who have attained some college or those who have attained some primary education. Using the middle three definitions for skilled labor, the difference $\rho - \theta$ is positive, and therefore consistent with capital–skill complementarity, but this difference is not significantly different from zero. Notice, however, that a likelihood ratio test of the restriction that $\rho = \theta$ is rejected ($p < .10$) only when the definition of skilled labor is set at one of the two lowest thresholds, either the attained primary or completed primary threshold.

When we use NLLS to estimate the nonlinear, two–level CES specification with fixed effects removed, specification (7) – see the "NLLS with FE" column in Table 1 – we find that there is evidence favoring capital–skill complementarity when skilled labor is defined using the lowest three thresholds: those who have attained some primary education, those who have completed primary education and those who have attained some secondary education. Defining skilled labor above the "attained secondary" threshold, evidence of capital–skill complementarity vanishes; it even appears that there is evidence of capital–skill substitutability, as the estimated difference $\rho - \theta$ becomes negative, though this difference is not found to be significant. The evidence for capital–skill complementarity using a low skill threshold appears to derive from the significantly positive estimates of ρ; the estimates of θ are never found to be significantly different from zero. Likelihood ratio tests confirm that the null hypothesis, $\rho = \theta$ can be rejected ($p < .10$) only when skilled labor is defined using one of the three lowest thresholds.

Applying the GMM–IV estimator to the first differenced two–level CES specification (7) yields positive estimates of the difference $\rho - \theta$ for all five definitions of skilled labor – see the "GMM–IV with FE" column in Table 1, suggesting capital–skill complementarity. However, the difference $\rho - \theta$ is never found to be significantly different from zero. Indeed, with just two exceptions, the

difference equation. They find that this alternative "systems approach" yields lower standard errors as compared with the GMM first–difference estimator of Arellano and Bond (1991) when applied to linear models. It is unclear whether the benefits of the systems estimator would extend to the nonlinear production function specification that we estimate. Furthermore, applying this approach would come at the cost of further reducing the number of observations we have available below the 219 observations we report for the GMM–IV estimator in Tables 1 and A3. We leave such an exercise to future research.

estimated values of ρ and θ are, by themselves, never significantly different from zero. While the GMM–IV coefficient estimates don't differ that much from the corresponding NLLS estimates, the standard errors of the GMM–IV estimates are 2 to 25 times higher than the corresponding NLLS estimates.

The statistical insignificance of parameter estimates using first difference GMM estimators is a well–known drawback of this technique; researchers estimating production functions with micro datasets have encountered the same difficulty – see the survey by Griliches and Mairesse (1998). As these two authors note, it seems that efforts to control for heterogeneity and endogeneity lead only to "exacerbations of other problems and misspecifications." (p. 198). One potential remedy, more (and better measured) data, will only come with the passage of time. Still, the GMM approach is preferred to NLLS estimation because the GMM approach represents an effort to address the endogeneity problem. While we may not currently have enough data, we *can* conduct a Monte Carlo exercise to assess how well our GMM–IV estimator fares (relative to NLLS) in detecting capital–skill complementarity in small sample sizes. This exercise is performed in section 6. To foreshadow our findings, the Monte Carlo test of the GMM–IV estimator yields estimates similar to those reported in Table 1 along with large standard errors. Hence our failure to detect capital–skill complementarity using the GMM–IV estimator may well be due to the small sample properties of the estimator and not to the absence of capital–skill complementarity in the available aggregate data.

Consider next, the CES–nested–in–CD specifications (10–11). Recall that these specifications are just restricted versions of the more general two–level CES specification (1).[18] In particular, the restriction is that in the more general specification (1), ρ is equal to zero, so that the elasticity of substitution between capital and unskilled labor $\sigma_{K,N}$ and (symmetrically) between unskilled and skilled labor, $\sigma_{N,S}$ are equal to unity. We can test this restriction by simply examining whether the estimates of ρ reported in Table 1 for the more general, two–level CES specification are significantly different from zero. For the NLLS and NLLS–FE estimates, ρ is *always* positive and significantly different from zero. The same holds true for the GMM–IV estimates of ρ when skilled labor is defined as those who have at least completed a secondary education. We are therefore able to reject the CES–nested–in–CD specification in these cases in favor of the more general two–level

[18]Given our findings for the two–level CES specification (2) we ignore the CES–nested–in–CD specifications (9) and (12) that are restricted versions of specification (2).

CES specification as the preferred specification.[19] Still, for completeness, we report NLLS and GMM–IV estimates of specifications (10–11) in Appendix 3.

While our main focus is on the presence or absence of capital–skill complementarity, our estimates of ρ in Table 1 are also of interest for the estimates they imply for the Allen elasticity of substitution between and unskilled labor and skilled labor (or capital) defined by $\sigma^a_{N,S} = 1/(1-\rho)$. The implied estimates of $\sigma^a_{N,S}$ using the estimates of ρ in Table 1, range from 1.3 to 10.[20] We note that Klenow and Rodríguez–Clare (1997) use aggregate, cross–country data to estimate the elasticity of substitution between workers with a primary education or less and workers with more education. Their estimate for this elasticity is 65! Thus our seemingly high estimates of the elasticity of substitution between unskilled and skilled labor when the skill threshold is low (e.g. completed primary education) are not without precedent.

5.3 Discussion of the Estimation Results

To summarize, our main finding is that using a time–series, cross section panel of 73 countries, there appears to be some evidence in support of the capital–skill complementarity hypothesis. The NLLS results from Table 1 are broadly supportive of the capital–skill complementarity hypothesis, though the evidence is often not statistically significant. However, when we address the endogeneity issue by using a GMM–IV estimator, the evidence for capital–skill complementarity is greatly weakened; the standard errors associated with the GMM–IV estimator do not allow us to make any real inference as to whether there is capital–skill complementarity or substitutability.[21] This finding is consistent with the possibility that over countries and across time, the extent of capital–skill complementarity (or substitutability) is subject to change, as argued by Goldin and Katz (1998). It may also be an artifact of our relatively small sample, and we will address this possibility later in the paper.

With regard to production function specifications, we argue against the use of the two–level specification (2) in favor of the two–level specification (1). Furthermore, we are able to reject the more restrictive CES–nested–in–CD specification, in favor of the more general, two–level specifica-

[19] Our rejection of the restricted CES–nested–in–CD specification is consistent with the findings of Krusell et al. (2000) who obtained the same finding using only U.S. data.

[20] In one case, the GMM–IV estimate of ρ when skilled labor is defined as those who have completed secondary education, the elasticity of substitution measure is ill–defined as the estimate of ρ exceeds unity.

[21] We note that our lack of strong evidence for capital–skill complementarity is consistent with the work of Caselli and Coleman (2002a) who obtain a similar result using a more indirect estimation method.

tion (1).

Finally, we note that our coefficient estimates shed some light on the appropriate definition of skilled labor, at least for purposes of assessing the issue of capital–skilled labor complementarity. The NLLS and NLLS with FE estimates in Table 1 suggest that a low threshold for dividing labor into unskilled and skilled classifications, such as whether workers have completed primary education, may be more conducive to a finding of capital–skill complementarity. This is a much lower threshold than has traditionally been considered in the literature (e.g. Krusell et al. (2000) define skilled workers as those who have completed a post–secondary (college) education or better).

5.4 Robustness of the Results using Adjusted Skilled Labor Data

We have examined the robustness of our results by considering an alternative and possibly more appropriate definition for skilled/unskilled labor. As discussed earlier, this "weighted" labor dataset adjusts for disparities in efficiency units across workers who belong to different educational subgroups within the class of workers we have designated as skilled or unskilled labor. Adjusting the measures of skilled and unskilled labor for the returns earned by the various educational subgroups provides us with a more precise measure of the contribution of skilled labor to output. Further details concerning the construction of this weighted labor data can be found in Appendix 1.

Unfortunately due to a lack of data on returns to schooling for all 73 countries, this adjustment to the labor data eliminates approximately one–third of the countries in our sample; we have 49 countries left, yielding just 294 observations (as compared with the 438 observations available in the full sample). Large sample sizes are particularly crucial to our work, as the results from estimating (the curvature of) the highly nonlinear nested CES production specifications requires a sufficiently large number of observations. Indeed, the GMM–IV estimation procedure for the nonlinear models, which requires the use of instruments, reduces the sample size even further to just 147 observations; the results from applying this procedure to the smaller weighted–labor dataset were unreliable resulting in economically implausible coefficient estimates and are not reported. The results from applying NLLS to the two–level model and the log–difference version of this model using the weighted–labor data (for which 294 observations were available) are presented in Tables A4 in Appendix 4.

The results in Table A4, the analog of Table 1, reveal that using the weighted labor data, the NLLS and NLLS–FE estimates of the difference $\rho - \theta$ are positive for most skilled labor definitions, but are never significantly different from zero. As for the CES–nested–in–CD specification, the restriction that $\rho = 0$ is clearly rejected based on the estimates reported in Table A4. However, for completeness, we report the CES–nested–in–CD estimates using the weighted labor data (the analog of Table A3) in table A5 of Appendix 4. Despite some differences, the estimation results using the weighted labor data are qualitatively similar to those we obtained using the unweighted–labor dataset. In particular, two of our main findings, weak evidence for capital–skill complementarity, and the rejection of the more restrictive CES–in–CD specification in favor of the general two–level specification continue to hold when we use the weighted labor data.

We have also tried to split the data to examine the sensitivity of our results to different subsamples of countries but to date, our estimates from such sample splits have been empirically implausible. We think this is due to having a limited number of observations that can not adequately capture variation in the curvature of our aggregate production functions.

6 Monte Carlo Experiments

Our main findings rest on the parameter estimates that we report in Table 1 (as well as Tables A3–A5). A natural question concerns the reliability of the estimates we have obtained using nonlinear estimation techniques for the two–level CES specification given our "small" samples and potential problems with fixed effects and endogeneity. Indeed, Kumar and Gapinski (1974) and Thursby (1980) report results from Monte Carlo experiments examining the small sample properties of CES parameter estimates obtained using nonlinear and linear estimation procedures and find that all of the CES parameter estimates were reliable with the notable exception of the elasticity of substitution parameter estimate! Since this estimate is the primary concern of our study, we felt it necessary to undertake our own Monte Carlo experiments, which we describe below. We note that Kumar and Gapinski and Thursby examined only the standard CES specification, not the two–level specification that we examine, and they focused on linear and nonlinear estimation techniques that differ from those used in this study. Furthermore, they used far fewer observations than we have available in our panel dataset (e.g. Thursby used just 20 observations) and their time series data had no cross–section component. Since we look at cross–country, time series data, our variables are likely to show (after controlling for the time trend) much more variation. For all of these reasons, a new set of Monte Carlo experiments seems warranted.[22]

The focus of our Monte Carlo experiments is on the small–sample properties and the potential problems of fixed–effects and endogeneity of the NLLS estimators of the two–level CES parameters, ρ and θ. In principle, we could examine the properties of the estimators we consider for all of the specifications suggested in the paper (using all the proxies for skilled labor and both the unweighted and weighted labor data). As our aim is the more limited one of providing some assessment of the reliability of the various estimators we employ, we have chosen to focus attention on the two–level CES specification (the most unrestricted nested CES specification) and to use only the unweighted data in our Monte Carlo experiments. As a proxy for skilled labor, we have chosen to use workers who have completed primary education ($S4$).

In particular, we consider the stochastic counterpart of specification (1) given by

$$\log Y_{it} = \log A_0 + \lambda t + \frac{1}{\rho} \log \left[a[bK_{it}^{\theta} + (1-b)S4_{it}^{\theta}]^{\rho/\theta} + (1-a)N4_{it}^{\rho} \right] + \varepsilon_{it}, \tag{13}$$

and

$$\log \left(\frac{Y_{it}}{Y_{i,t-1}} \right) = \lambda + \frac{1}{\rho} \log \frac{\left[a[bK_{it}^{\theta} + (1-b)S4_{it}^{\theta}]^{\rho/\theta} + (1-a)N4_{it}^{\rho} \right]}{\left[a[bK_{i,t-1}^{\theta} + (1-b)S4_{i,t-1}^{\theta}]^{\rho/\theta} + (1-a)N4_{i,t-1}^{\rho} \right]} + \varepsilon_{it} - \varepsilon_{i,t-1}, \tag{14}$$

where ε_{it} is a disturbance term.

A critical question concerning our simulation exercise is how to generate disturbances with built–in endogeneity and fixed–effects problems.[23] Without any prior knowledge of the magnitude of the endogeneity problem inherent in (13–14) nor the way by which it affects ε_{it}, we construct a disturbance term in the following simple but intuitive way:

$$\varepsilon_{it} = \bar{\varepsilon}_{it} + \alpha \left(\eta_{K,it} + \eta_{S,it} + \eta_{N,it} \right) + \beta \left(\bar{\eta}_{K,i} + \bar{\eta}_{S,i} + \bar{\eta}_{N,i} \right). \tag{15}$$

In equation (15), $\bar{\varepsilon}_{it}$ is a randomly generated *i.i.d.* component that is orthogonal to the explanatory variables, and the terms in the transposed vector $(\eta_{K,it}, \eta_{S,it}, \eta_{N,it}) = \eta'_{it}$ are the disturbances for country i at date t from the regression

$$X_{it} = \Gamma Z_{i,t-2} + \eta_{it}, \tag{16}$$

[22] To our knowledge, there is no prior work examining the small sample properties of estimates obtained from nonlinear or linear estimation of the two–level CES specification that we consider in this paper. Thus our Monte Carlo experiments are of independent interest beyond our application examining the capital–skill complementarity hypothesis.

[23] The authors are grateful to an anonymous referee's suggestion for generating errors with built-in endogeneity.

where $X'_{it} = (K_{it}, S_{it}, N_{it})$, $Z'_{it} = (X'_{it}, Y_{it})$, and Γ is a 3×4 matrix of estimated parameters. By construction, the last equality says that the disturbance vector η_{it} is orthogonal to $Z_{i,t-2}$, but is correlated with X_{it}.[24] Given the relationship between η_{it} and ε_t proposed in equation (15), this has two main implications: (i) the component $(\eta_{K,it} + \eta_{S,it} + \eta_{N,it})$ in equation (15) introduces an endogeneity problem because it means that $corr(X_{it}, \varepsilon_t) \neq 0$; and (ii) $corr(Z_{it-2}, \varepsilon_t) = 0$, that is, the second lags of the explanatory and dependent variables represent good instruments in order to estimate specification (14).[25] On the other hand, the country–specific averages $\bar{\eta}_{K,i}$, $\bar{\eta}_{S,i}$ and $\bar{\eta}_{N,i}$ in expression (15), with $\bar{\eta}_{J,i} = (1/T)\sum_{t=1}^{T} \eta_{J,it}$, incorporate fixed effects. Thus our disturbance term, ε_{it}, is comprised of both endogeneity and fixed effect components. Obviously, the larger are the parameters α and β in absolute value, the larger is the variance of these components and, consequently, the relative size of the endogeneity and fixed effect problems built into the disturbance term.

Once the disturbance terms are defined, (13) and (14) are used to generate data on output, Y, employing our panel data of 73 countries over six 5–year–interval periods for given values of capital, K, unweighted skilled labor, $S4$, and unskilled labor, $N4$. In all Monte Carlo experiments, the four parameters of the production functions were always set as follows: $A_0 = 1$, $\lambda = 0.02$, $a = 0.4$, $b = 0.5$; these values fall within the range of coefficient estimates we obtained from our NLLS regressions. We chose the elasticity of substitution parameters $\rho = 0.3$ and $\theta = 0.1$ to allow for capital–skill complementarity (i.e. $\rho - \theta = 0.2 > 0$).

Another important consideration is the choice of the variance for the random disturbances. Large values for $\sigma^2_{\varepsilon_{it}}$ and $\sigma^2_{\varepsilon_{it}-\varepsilon_{i,t-1}}$ would yield output series from specifications (13–14) that were almost purely random. By contrast, very small values for $\sigma^2_{\varepsilon_{it}}$ and $\sigma^2_{\varepsilon_{it}-\varepsilon_{i,t-1}}$ would result in output series that were nearly deterministic. The variances for equations (13–14) were chosen according to the rule used in Kumar and Gapinski (1974) and Thursby (1980): the variances were chosen to yield certain R^2 values for the NLLS regressions. In particular, the rule used to obtain these

[24] In the Monte Carlo experiments, we obtained $corr(X_{it}, \eta_{it}) \in [0.07, 0.10]$, $|\,corr(Z_{it-2}, \eta_{it})\,| < 10^{-12}$.

[25] The correlation between the explanatory variables, both in levels and first differences, and the instruments is high. In particular, the average value across the different variables contained in vectors X_t and Z_t equals

$$corr(X_t, Z_{t-2}) = corr(X_{t-1}, Z_{t-2}) = corr(X_t, Z_{t-3}) = corr(X_{t-1}, Z_{t-3}) = 0.69,$$

and

$$corr(X_t - X_{t-1}, Z_{t-2}) = corr(X_t - X_{t-1}, Z_{t-3}) = 0.56.$$

variances is

$$\sigma^2_{\varepsilon_{it}} = var(\log Y_t)(1 - R^2_{\varepsilon_{it}}), \tag{17}$$

$$\sigma^2_{\varepsilon_{it}-\varepsilon_{i,t-1}} = var[\log(Y_t/Y_{t-1})](1 - R^2_{\varepsilon_{it}-\varepsilon_{i,t-1}}), \tag{18}$$

where $R^2_{\varepsilon_{it}} = 0.96$, and $R^2_{\varepsilon_{it}-\varepsilon_{i,t-1}} = 0.25$. These R^2 values were obtained from NLLS regression and differenced NLLS estimation, respectively. Thus we chose $\sigma^2_{\varepsilon_{it}} = 0.11676$ and $\sigma^2_{\varepsilon_{it}-\varepsilon_{i,t-1}} = 0.01659$.

For each of the 73 countries, we have 6 annual observations. However, given that we use the second lag of the production function variables to construct artificial disturbances, we were restricted to generating for each trial of the Monte Carlo experiment only 292 ($(6-2) \times 73$) observations on ε_{it} using a random number generator. A total of 100 sets of 292 ε_{it} values were constructed in this fashion. Using these 100 disturbance sets we built 100 corresponding sets of artificial output data (Y) using the actual data on capital, skilled and unskilled labor, and holding constant our parameter choices for the two–level CES function, ρ, θ, A_0, λ, a, b. For the NLLS estimation employing these simulated data, the true parameter values were used as initial guesses in the hope such choices would minimize the number of iterations required for convergence.[26]

We performed four Monte Carlo experiments. In the first one, we assume that there is neither fixed effect nor endogeneity problems and set $\alpha = \beta = 0$; that is, the disturbance term is completely orthogonal, with variance equal to $\sigma^2_{\varepsilon_{it}} = 0.11676$ as discussed above. The second experiment introduces fixed effects (i.e. $\alpha = 0, \beta > 0$). In this case, we choose a value of β so that the variance of the differenced errors is $\sigma^2_{\varepsilon_{it}-\varepsilon_{i,t-1}} = 0.01659$, as discussed previously. The third and fourth exercises use disturbances that incorporate both fixed effect and endogeneity problems (i.e. $\beta, \alpha > 0$). We maintain the value of β used in the second experiment, and to assign a number to α we use the fact that equation (15) implies that the variance of the differenced errors can be decomposed as

$$\begin{aligned} \sigma^2_{\varepsilon_{it}-\varepsilon_{it-1}} &= var\left(\bar{\varepsilon}_{it} - \bar{\varepsilon}_{it-1}\right) \\ &\quad + \alpha^2 var\left([\eta_{K,it} - \eta_{K,it-1}] + [\eta_{S4,it} - \eta_{S4,it-1}] + [\eta_{N4,it} - \eta_{N4,it-1}]\right). \end{aligned} \tag{19}$$

In the third experiment, we suppose that the first summand on the right hand side of expression

[26] In some cases, especially with GMM estimation, the algorithm did not converge when the estimates of the parameters a and b seemed to be very close to either 0 or 1. In these cases, we assigned a value of either 0.01 (if close to 0) or 0.99 (if close to 1) to the coefficient that was causing the problem, and estimated the remaining parameters.

(19), $var(\bar{\varepsilon}_{it} - \bar{\varepsilon}_{it-1})$, equals the second right hand side term. In other words, the contributions of the orthogonal component and the endogenous component to the variance of the differenced errors are assumed to be equal. In the fourth experiment, we amplify the importance of the endogeneity problem, and assume that only 25% of the variance in the difference errors comes from the orthogonal component, i.e. $var(\bar{\varepsilon}_{it} - \bar{\varepsilon}_{it-1}) = (1/4)\sigma^2_{\varepsilon_{it}-\varepsilon_{it-1}}$, while the endogenous component contributes the remaining 75% of the variance. In each experiment, we apply the three estimation techniques used previously: (i) NLLS on levels (292 observations); (ii) NLLS on the first–differenced specification, which corrects for fixed effects (219 observations); and (iii) GMM on the first–differenced specification employing 8 instruments, the $t-2$ and $t-3$ dated lags of Y, K, $S4$, and $N4$ as was our practice using the GMM–IV estimator (219 observations). The results are reported in Panels A–C of Table 2. Each panel refers to a different estimation technique and reports the results from applying that technique to the four different error scenarios. Thus we can see how the performance of the estimator evolves as we change the error structure. In each case, the table provides the sample mean, standard deviation and bias of the estimates of ρ and θ. The last column of Table 2 reports the root mean squared error (RMSE) to facilitate comparisons across the different estimators.[27]

As shown in Panel A, when errors are orthogonal the sample mean NLLS estimates of ρ and θ are relatively close to their true values (0.3 and 0.1 respectively) and have small biases (2.5% and 3.4%, respectively). However, the estimates are imprecise. This is especially true for estimates of the parameter θ which had a standard deviation of 0.2629, more than twice the mean estimate. When the error incorporates fixed effects, the precision of the estimates rises because the stochastic part that varies from sample to sample is exclusively due to the orthogonal component that has a lower weight when other components are included. Indeed, in the second row of Panel A we see that the dispersion measures of the estimates of ρ and θ fall to 0.0142 and 0.0258, respectively. Notice, however, that the fixed effect problem causes an important bias that amounts to -0.0509 (i.e., 17%) for ρ and 0.7210 (above 700%) for θ. The big deterioration in the estimation of θ when fixed effects are introduced is clear if we look at the RMSE; its value more than triples. The third and fourth rows of Panel A present the results when the error term is composed of both fixed effect and endogeneity components. The precision increases with the contribution of the endogeneity component for the reasons given above. In addition, the simultaneity problem induces an upward bias that, along with the precision of the estimates, increases with the contribution

[27]RMSE is defined as the squared root of the sum of the variance and the squared bias.

Table 2: Monte Carlo Experiments

Panel A: NLLS

Errors	Parameter	Mean	St. Dev.	Bias	RMSE
Orthogonal	ρ	0.3076	0.0985	0.0076	0.0990
	θ	0.0966	0.2629	-0.0034	0.2629
With fixed effects (FE)	ρ	0.2491	0.0142	-0.0509	0.0529
	θ	0.8210	0.0258	0.7210	0.7215
With FE and endogeneity (50% – 50% contribution)	ρ	0.2633	0.0108	-0.0367	0.0387
	θ	0.8353	0.0252	0.7353	0.7357
With FE and endogeneity (25% – 75% contribution)	ρ	0.2655	0.0071	-0.0345	0.0346
	θ	0.8432	0.0169	0.7432	0.7434

Panel B: NLLS with FE

Errors	Parameter	Mean	St. Dev.	Bias	RMSE
Orthogonal	ρ	0.4151	0.4449	0.1151	0.4595
	θ	0.1160	0.4759	0.0160	0.4762
With fixed effects (FE)	ρ	0.3018	0.1011	0.0018	0.1010
	θ	0.0943	0.2385	-0.0057	0.2385
With FE and endogeneity (50% – 50% contribution)	ρ	0.4192	0.0728	0.1192	0.1396
	θ	0.5069	0.1307	0.4069	0.4274
With FE and endogeneity (25% – 75% contribution)	ρ	0.4513	0.0476	0.1513	0.1584
	θ	0.5709	0.0957	0.4709	0.4805

Panel C: GMM–IV with FE (8 instruments)

Errors	Parameter	Mean	St. Dev.	Bias	RMSE
Orthogonal	ρ	0.9279	0.9478	0.6279	1.1369
	θ	0.5409	0.9020	0.4409	1.0039
With fixed effects (FE)	ρ	0.6767	0.7905	0.3767	0.8756
	θ	0.3468	0.6693	0.2468	0.7133
With FE and endogeneity (50% – 50% contribution)	ρ	0.3225	0.2139	0.0225	0.2152
	θ	0.1517	0.7914	0.0517	0.7932
With FE and endogeneity (25% – 75% contribution)	ρ	0.2962	0.2061	-0.0038	0.2062
	θ	-0.0979	0.7543	-0.1979	0.7798

Notes: In the Monte Carlo experiments we use the two–level CES specification 1 with 'completed primary' (S4) as the threshold for skilled labor and $\rho = 0.3, \theta = 0.1$.

of the endogeneity component. As a consequence, the mean NLLS estimate of ρ gets closer to its population value, whereas the mean NLLS estimate for θ rises further above its true value. Consequently, one would be led to conclude against the capital–skill complementarity hypothesis, as $\rho < \theta$, and the estimates appear to be quite precise. Overall, Panel A suggests that in small samples NLLS estimation delivers consistent estimates only when errors are completely orthogonal. Moreover, in all cases, the dispersion is relatively high, thus making the estimates fairly imprecise especially for the parameter θ.[28]

Panel B presents the results of performing NLLS estimation on the first–difference specification. The use of this specification results in two changes with respect to the previous estimation method. First, by differencing the data we lose another 73 observations, which further weakens the estimation precision. Second, the first–difference model induces serial correlation among consecutive errors because the disturbances now follow a MA(1) process, $\varepsilon_{it-}\varepsilon_{i,t-1}$. These two added features are probably responsible for the worse outcome we observe when errors are orthogonal. In particular, comparing the first row of Panel B to the first row of Panel A, we observe much more imprecise and biased estimates in the former panel, showing a RMSE that is 4.6 and 1.8 larger for ρ and θ, respectively. However, when we estimate using artificial samples that incorporate fixed effects, the benefit from using the first–difference estimator becomes evident. The second row of Panel B shows that correcting for fixed effect problems provides mean NLLS estimates that are quite close to their population values with a bias of 0.6% and 5.7%, respectively. Similar to the NLLS estimates in Panel A, precision is low, with a relatively large standard deviation for estimates of θ. The introduction of endogeneity problems generates an upward bias in the estimates and, as in Panel A, is also associated with a small improvement in precision. Both the bias and the improvement in precision increase as the endogeneity component becomes greater, causing the overall performance of the first–difference estimator to worsen. For example, when the endogeneity component accounts for 75% of the differenced–error variance, the RMSEs for ρ and θ are more than 50% above the values obtained when the error term is composed of only fixed effect and orthogonal components. The conclusion from Panel B is that, in small samples, NLLS estimation of the first–difference model delivers consistent estimates only when errors incorporate both fixed effect and orthogonal

[28]We also performed a Monte Carlo experiment for NLLS estimation using 438 observations, as we had in the previous section. The standard deviations of the estimates were lower (as expected), but still substantial. For example, in the orthogonal case, the standard deviations equal 0.0688 for ρ and 0.1890 for θ.

components. In all cases, estimation is fairly imprecise, especially for the parameter θ.[29]

Estimates from applying GMM to the first–difference model are presented in Panel C. The first striking result is that dispersion rises enormously. For example, in the case of orthogonal errors, the estimates for ρ and θ have standard deviations of 0.9478 and 0.9020, respectively, and a substantial bias of 0.6279 (200%) for ρ and 0.4409 (above 400%) for θ. Both the dispersion and the bias decline as weight is given to fixed effect and endogeneity components. As in Panels A and B, the dispersion decreases with the decline in the variance of the orthogonal part. The bias, in turn, declines as well, as GMM appears to be working to correct (at least partially) the simultaneity problem. The improvement is evident as we go down the rows of Panel C. The bias of the mean parameter estimate of ρ declines to -0.0038 (1.26%) when endogeneity explains 75% of the differenced–error variance, whereas for estimates of θ the bias obtains its minimum value of -0.0517 (51.7%) when simultaneity explains 50% of the differenced–error variance. Notice that these values are the minimum biases found across all three estimation techniques when endogeneity is present. However, taking into account the large dispersion delivered by GMM, its performance must still be regarded as the worst of three estimators. This case can be made by comparing the values of the RMSE across the three panels: except in the fixed–effect–error case for the parameter θ, GMM estimation yields the highest RMSE values over all experiments.[30]

Summarizing this section, our Monte Carlo experiments suggest that NLLS estimation of the two–level CES specification (with and without correcting for fixed effects) provides fairly consistent estimates of the elasticity of substitution parameters, ρ and θ, when it should. GMM estimation, on the other hand, obtains highly consistent estimates only for ρ when it should. However, all techniques provide imprecise estimates. The problem seems to be especially serious for estimates of the parameter θ, which are substantially upward biased in the presence of both fixed effects and endogeneity problems. Indeed, if fixed effects and endogeneity problems in the data were of the same form and incidence as those found in our Monte Carlo study, the biased estimates of θ might lead us to erroneously reject the null hypothesis of no capital–skill complementarity. We conclude that small–sample estimates using NLLS–level, NLLS–FE, and GMM–IV of the elasticity

[29]A Monte Carlo experiment for NLLS estimation on first–differenced data using 365 observations (the number that we had in the previous section) delivered standard deviations equal to 0.0703 for ρ and 0.1541 for θ in the best scenario (i.e., with errors composed of just fixed–effect and orthogonality components).

[30]The poor performance of the GMM estimator in the small sample size of our experiments is consistent with findings from several other recent studies that seek to identify the parameters of production functions using small micro–level panel data sets, as surveyed in Griliches and Mairesse (1998).

of substitution parameters in the two–level CES specification should be taken with caution given that their reliability appears to be relatively low.

7 Conclusions

The aim of this paper is to examine the cross–country evidence for capital–skilled labor complementarity using aggregate production function specifications and a time–series, cross–section panel of countries. In particular, we address three empirical questions. First, what is the best specification of the aggregate production technology for purposes of examining the capital–skill complementarity hypothesis? Second, how do we define skilled labor. Finally, is there any cross–country evidence in support of the capital–skill complementarity hypothesis? With regard to the first issue we argue that specification (1) is preferred to specification (2). Furthermore, within specification (1) we find that we can reject the restricted, CES–nested–in–CD specification in favor of the more general two–level CES form. Second, unlike other empirical studies, we consider five different methods of classifying skilled labor as the appropriate threshold for dividing workers into skilled and unskilled classes in a cross–country study is not at all clear. Finally, and perhaps most importantly, we find some evidence in support of the capital–skill complementarity hypothesis. The case for capital–skill complementarity can be made using our NLLS estimates of the first differenced model, which takes account of country specific fixed effects. Using those estimates, as reported in Table 1, evidence for capital–skill complementarity obtains when the threshold for defining a skilled worker is very low, e.g. workers who have attained *some* primary education, or who have completed primary or attained some secondary education. These thresholds are all lower than the completed secondary, or post–secondary education threshold for skilled labor that is more typically encountered in the literature.

Still, we urge caution in taking these NLLS estimates for the differenced model too seriously. First, we found that when the data within the skilled and unskilled classifications for labor are adjusted for returns to schooling, the NLLS–FE estimates supporting capital–skill complementarity cease to be statistically significant, as shown in Table A4. Second, our NLLS first–difference estimator fares poorly in our Monte Carlo study when both fixed effect and endogeneity problems are present. Finally, and most importantly, our preferred estimation technique, the GMM–IV first difference estimator which addresses the endogeneity problem, yields biased estimates, especially for θ which might erroneously lead to us to reject the null hypothesis of no capital–skill complementarity. Another problem with the GMM–IV estimator is the large standard errors associated with the estimates.

We conclude that there is some evidence in support of the capital–skill complementarity hypothesis at the aggregate production level, however the evidence is not very strong. Setting aside the data and estimation problems, an intriguing alternative explanation for our weak evidence is that the extent of capital–skill complementarity (or substitutability) varies with a country's stage of development and is therefore subject to change over time, as Goldin and Katz (1998) have convincingly argued. If this hypothesis is true, then, consistent with our findings, evidence in support of the capital–skill complementarity hypothesis should be especially difficult to obtain using a time–series, cross–section panel of countries. Finally, our findings have implications for the debate concerning the source of rising wage and income inequality across countries. Some authors, e.g. Krusell et al. (2000) have pointed to capital–skill complementarity as the likely source of this phenomenon. Our lack of strong evidence for capital–skill complementarity suggests that researchers might want to consider alternative, complementary explanations for rising inequality, for example, skill–biased technical change.

Appendix 1: Data

The data used in this paper (unweighted and weighted) are available from the authors upon request.

- *Income (Y)* [Source: PWT–5.6]

Cross–country real GDP per worker and real GDP per capita are in constant dollars (1985 international prices) using the Chain index as described by Summers and Heston (1991). These data are from the Penn World Tables (PWT), Version 5.6 and are available on–line at:
http://datacentre.chass.utoronto.ca/pwt/index.html.

- *Physical capital stocks (K)* [Source: PWT–5.6]

Physical capital is constructed using the perpetual inventory approach with investment shares data obtained from PWT–5.6. In particular, the physical capital stock is calculated by summing investment from its earliest available year (1960 or earlier) to 1990 with the annual depreciation rate fixed at 6 percent. The initial physical capital stock is determined by the initial investment rate, divided by the depreciation rate plus the growth rate of investment during the subsequent ten years. See Duffy and Papageorgiou (2000) for further details concerning this procedure.

- *Skilled and Unskilled Labor (S, N)* [Source: Barro and Lee (2001), and Lee (2001)]

We construct five alternative proxies for skilled and unskilled labor as the definition of skilled/unskilled labor is arbitrary. These proxies are constructed using achievement rate data from Barro and Lee (2001) and multiplying these rates by the sized of the total labor force. Our five proxies for skilled and unskilled labor are as follows:

Unweighted data

1. $S1$ is equal to the number of workers that have attained at least some post–secondary education and $N1$ is equal to the rest of the workers in the labor force.
2. $S2$ is equal to the number of workers that have completed secondary education, and $N2$ is equal to the rest of the workers in the labor force.
3. $S3$ is equal to the number of workers that have attained at least some secondary education, and $N3$ is equal to the rest of the workers in the labor force.
4. $S4$ is equal to the number of workers that have completed primary education, and $N4$ is equal to the rest of the workers in the labor force.
5. $S5$ is equal to the number of workers that have attained at least some primary education, and $N5$ is equal to the rest of the workers in the labor force.

Weighted data

Within a given skill class say, Si or Ni, $i = 1,2,3,4$ or 5, we weigh individuals by a function of the length in years of their schooling level times the return to schooling. In addition, the aggregate value is constructed so that it is measured in terms of the efficiency units of the lowest educational subcategory included in the skill class. Lengths of educational attainments subgroups by country are from Lee (2001). Returns to schooling by nation are taken from Bils and Klenow (2000), and were obtained following the Mincerian approach which assumes that log–wages are linear in years of schooling.

An example: Let $l_{i,j}$ be the length in years of educational level j in country i, $L_{i,j}$ the number of workers with this schooling level, and ϕ_i is the Mincerian return in country i. For nation i, $S2$ and $N2$ are computed as follows:

$$\begin{aligned} S2(i) &= L_{i,cs} + \exp(\phi_i \, l_{i,sps}) \, L_{i,sps} + \exp(\phi_i \, l_{i,cps}) \, L_{i,cps}, \\ N2(i) &= L_{i,up} + \exp(\phi_i \, l_{i,ap}) \, L_{i,ap} + \exp(\phi_i \, l_{i,cp}) \, L_{i,cp} + \exp[\phi_i \ (l_{i,cp} + l_{i,ss})] \ L_{i,ss}, \end{aligned}$$

where up, ap, cp, ss, cs, sps and cps denote uneducated population, attained primary, completed primary, some secondary, completed secondary, some post–secondary and completed post–secondary education, respectively.
The Barro and Lee (2001) dataset is available on–line at: http://www2.cid.harvard.edu/ciddata.

- *Labor Force* [Source: PWT–5.6]

The cross–country dataset on the labor force is calculated from the PWT–5.6 series on GDP per capita and GDP per worker. It represents the population between the ages of 15 and 65 (taken to represent the labor force).

Table A1: Mean Values of Unweighted–Data from the 73 Country Sample

Country	Code	GDP (mill. US$)	Capital (mill. US$)	Skilled Lab. ($S4$) (thous.)	Unskilled Lab. ($N4$) (thous.)
Algeria	DZA	44620.1	97958.8	954.0	3022.5
Argentina	ARG	150309.3	285059.0	5911.3	4144.3
Australia	AUS	173178.0	522017.9	5654.5	757.6
Austria	AUT	71539.6	188862.3	2825.5	543.1
Bangladesh	BGD	98086.8	42154.0	4178.7	20201.8
Brazil	BRA	406421.4	733594.1	9889.7	30922.6
Belgium	BEL	98489.2	259548.9	3041.6	818.4
Bolivia	BOL	9241.7	17061.4	673.3	1030.7
Canada	CAN	308801.9	758688.9	9386.5	1313.7
Chile	CHL	39315.5	85170.8	1848.9	1818.6
Colombia	COL	68509.0	108933.2	2841.5	5067.1
Costa Rica	CRI	7094.8	6146.6	294.0	433.6
Cyprus	CYP	3233.0	8625.1	188.1	94.5
Denmark	DEN	55921.8	160162.7	2174.9	412.1
Ecuador	ECU	20081.0	41964.5	1108.1	1257.0
El Salvador	SLV	7877.6	3517.4	312.8	1070.6
Finland	FIN	48897.2	172168.6	1675.2	669.2
France	FRA	576919.9	1625109.1	15998.7	7304.3
Germany	DEU	677584.3	1444383.0	23376.9	4322.1
Ghana	GHA	9813.9	7368.2	1217.3	3147.2
Greece	GRC	49610.5	118409.6	2588.5	1014.9
Guatemala	GTM	14366.8	14185.8	363.4	1571.4
Haiti	HTI	4739.6	2802.3	344.9	2044.3
Honduras	HND	4736.3	6786.2	257.8	801.6
Iceland	ICE	2245.1	6219.7	79.0	29.7
India	IND	631421.3	828804.9	60737.7	199061.9
Indonesia	IDN	180966.6	259253.0	17792.7	35987.6
Iran	IRN	153674.3	226936.6	2904.2	8085.5
Iraq	IRQ	62576.6	79507.4	694.7	2604.7
Ireland	IRL	21031.5	54008.1	921.3	308.6
Israel	ISR	27462.0	63439.3	967.7	394.7
Italy	ITA	513760.7	1453670.2	14399.5	7471.7
Jamaica	JAM	5086.1	14084.8	405.8	489.3
Japan	JPN	1085463.9	3199481.3	54701.0	15772.3
Jordan	JOR	6094.8	7796.0	222.0	337.3
Kenya	KEN	12896.4	22662.3	1452.4	5557.1

Notes: The sources for these data are PWT–5.6 and Barro and Lee (2001). Country specific mean values presented above have been rounded to the first decimal place.

Table A1: Mean Values of Unweighted–Data from the 73 Country Sample, continued.

Country	Code	GDP (mill. US$)	Capital (mill. US$)	Skilled Lab. ($S4$) (thous.)	Unskilled Lab. ($N4$) (thous.)
Korea, Rep.	KOR	123619.8	173122.2	10095.1	3834.1
Malawi	MWI	2970.0	2888.4	337.3	2351.9
Malaysia	MYS	46709.4	88587.3	2384.7	2654.9
Mali	MLI	3156.0	1892.7	89.1	2237.1
Mauritius	MUS	3660.0	3999.2	210.9	277.6
Mexico	MEX	325533.8	499518.5	8706.5	11274.1
Mozambique	MOZ	11780.6	2778.0	464.6	5938.5
Myanmar (Burma)	MMR	16679.8	14309.7	3230.0	11247.9
Netherlands	NLD	145453.1	384517.7	4196.1	1102.4
New Zealand	NZL	31867.4	57557.5	1120.8	150.5
Norway	NOR	44634.4	147087.0	1467.6	377.3
Pakistan	PAK	90718.0	79640.2	5266.0	18308.3
Panama	PAN	5486.0	10919.8	357.3	282.2
Paraguay	PRY	5844.9	7146.3	410.6	591.3
Peru	PER	43241.4	85691.6	2377.3	2677.2
Philippines	PHI	74413.5	115076.7	10196.4	6751.9
Portugal	PRT	44167.8	94405.4	1863.8	2162.5
Senegal	SEN	6137.7	4270.0	373.1	2127.5
Sierra Leone	SLE	3471.8	315.328.1	131.4	1108.3
Singapore	SGP	14973.3	36424.8	489.1	467.7
Spain	ESP	257028.1	637275.6	8356.0	4465.3
Sri Lanka	LKA	23021.0	12767.7	2794.0	2293.5
Sudan	SDN	14658.3	20107.2	599.6	5412.2
Sweden	SWE	99908.5	261627.0	2916.9	1085.6
Switzerland	CHE	87821.5	275816.1	2432.8	645.5
Tanzania	TZA	8161.8	8599.8	1068.1	7670.5
Thailand	THA	98267.9	137663.4	7447.0	14842.4
Tunisia	TUN	14045.3	17003.8	455.9	1372.7
Turkey	TUR	123388.5	238604.3	8362.6	10639.4
Uganda	UGA	6959.6	2171.8	791.5	4915.8
United Kingdom	GRB	563966.7	1132350.1	19371.8	7305.7
United States	USA	3307524.9	8438179.1	96343.5	6353.7
Uruguay	URY	12456.0	23513.3	606.4	528.4
Venezuela	VEN	95991.3	205740.7	1704.6	2806.1
Zaire	ZAR	13408.9	6921.8	1738.2	8877.0
Zambia	ZMB	5199.2	21789.5	434.9	1445.4
Zimbabwe	ZWE	8043.5	18997.5	734.3	2157.4

Notes: The sources for these data are PWT–5.6 and Barro and Lee (2001). Country specific mean values presented above have been rounded to the first decimal place.

Appendix 2: Estimation Results using Specification (2)

In this appendix we report estimates for the NLLS, NLLS with FE and GMM–IV estimators applied to specification (2). These estimates, as shown in Table A2, contrast sharply with the comparable estimates for specification (1) as reported in Table 1 in the text. In particular we note that with only a few exceptions, the estimates of ρ exceed unity, so that the Allen elasticity of substitution between unskilled and skilled labor (or capital), $\sigma^a_{N,S} = 1/(1-\rho)$, is ill–defined. We note also that there is large variation in the parameter estimates across the different estimation methods and across the different definitions for skilled labor. Finally, we observe that in some cases the standard errors are extremely large. Because the estimates reported in Table 1 for specification (1) are not generally prone to these same problems, we prefer the estimation results based on specification (1). We reached the same conclusion for the restricted, CES–nested–in–CD model.

Table A2: Estimates of ρ and θ using Specification 2

Skilled Labor	Parameter	NLLS	NLLS with FE	GMM–IV with FE
Attained College	ρ	5.2014 (4725017.1)	2.4443 (1.6622)	2.8675 (3.6370)
	θ	0.48122*** (.06274)	0.25496*** (0.07460)	0.35013 (1.3939)
	$\rho-\theta$	4.7202 (4725017.1)	2.1893 (1.6550)	2.5174 (2.4616)
Completed Secondary	ρ	7.3792 (9425088)	1.4185*** (0.43693)	2.9277 (6.5447)
	θ	0.49136*** (0.05967)	0.28844*** (0.07479)	1.0611*** (0.35565)
	$\rho-\theta$	6.8878 (9425088)	1.1301** (0.44429)	1.8666 (6.6016)
Attained Secondary	ρ	0.14967 (0.22456)	0.37839*** (0.07248)	2.6257 (5.3128)
	θ	0.60182*** (0.07542)	0.50824* (0.29852)	0.71885 (1.0334)
	$\rho-\theta$	−0.45215* (0.23468)	−012985 (0.30483)	1.9068 (6.1040)
Completed Primary	ρ	1.2403*** (0.36102)	1.4425*** (0.24337)	1.6674 (4.5032)
	θ	0.57290*** (0.06113)	−0.29092 (0.94936)	1.3498 (1.4783)
	$\rho-\theta$	0.66744* (0.07231)	1.7335 (0.86283)	0.31757 (4.7191)
Attained Primary	ρ	−0.13134 (0.11186)	0.48746** (0.19273)	2.9339 (11.892)
	θ	1.6009 (10985)	0.52590*** (0.08331)	0.48440* (0.27412)
	$\rho-\theta$	−1.7323 (10985)	−0.03843 (0.21586)	2.4495 (12.054)
Obs.		438	365	219

Notes: Standard errors, in parentheses, were recovered using standard approximation methods. White's heteroskedasticity correction was used. ***, **, * Significantly different from 0 at the 1% , 5%, and 10% levels, respectively.

Appendix 3: Estimates from the CES–Nested–in–CD Specification

Table A3 reports the estimates we obtained for versions of the nonlinear CES–nested–in–CD specification (10–11). Recall that for these specifications, capital–skill complementarity obtains if the estimated value of $\theta < 0$; estimates of $0 < \theta \leq 1$ imply capital–skilled labor *substitutability* and capital–*unskilled* labor complementarity. As Table A3 reveals, for the nonlinear CES–nested–in–CD specification, we do observe estimates of θ that are positive and significantly different from zero, implying capital–skilled labor substitutability. However, we note that the positive and highly significant NLLS estimates for θ are mainly observed in specifications that do not account for fixed effects or make use of instruments; in the fixed effects (first difference) specification without or with instruments (NLLS with FE) and (GMM–IV with FE), the estimates of θ are (with one exception) positive and (again with one exception) not significantly different from zero.

Table A3: CES–Nested–in–CD Nonlinear Estimation

Skilled Labor	Parameter	NLLS	NLLS with FE	GMM–IV with FE
Attained	θ	0.53668***	0.62413	0.08865
		(0.18897)	(0.50871)	(4.5838)
College	γ	0.64516***	0.43305***	0.15555
		(0.01179)	(0.03515)	(0.22575)
Completed	θ	0.31833*	0.33990	0.08018
		(0.16877)	(0.71577)	(2.8534)
Secondary	γ	0.70372***	0.63800	0.34563
		(0.01204)	(0.90695)	(0.21084)
Attained	θ	0.36866***	0.75042	0.02339
		(0.13689)	(0.68256)	(1.5102)
Secondary	γ	0.77621***	0.45653***	0.15443
		(0.01294)	(0.10581)	(0.48466)
Completed	θ	0.81919***	0.24676**	-0.04599
		(0.18284)	(0.12920)	(0.32467)
Primary	γ	0.83490***	0.82146***	0.03453
		(0.01458)	(0.02735)	(0.24487)
Attained	θ	0.94752***	0.94713***	0.13742
		(0.17444)	(0.01779)	(8.8275)
Primary	γ	0.89931***	0.31991***	0.31657
		(0.01130)	(0.10735)	(0.25896)
No. Obs.		438	365	219

Notes: Standard errors, in parentheses, were recovered using standard approximation methods. White's heteroskedasticity correction was used. ***, **, * Significantly different from 0 at the 1% , 5%, and 10% levels, respectively.

Appendix 4: Estimation Results using the Weighted–Labor Data

Here we report estimates for the specifications reported in Tables 1 and A3 using the weighted labor data as described in section 5.4.

Table A4: Two–Level CES Nonlinear Estimation (weighted–labor data)

Skilled Labor	Parameter	NLLS	NLLS with FE
Attained College	ρ	0.47987*** (0.11082)	0.36999*** (0.09621)
	θ	0.47505 (0.33170)	−0.55621 (1.4671)
	$\rho - \theta$	0.00481 (0.35710)	0.92621 (1.4646)
	LR Test $\rho = \theta$	0.00022	0.01924
	$p > \chi^2(df = 1)$	0.98796	0.63854
Completed Secondary	ρ	0.26689*** (0.10102)	0.45315*** (0.09682)
	θ	1.13733*** (0.31369)	0.84715** (0.41589)
	$\rho - \theta$	−0.87045*** (0.33146)	−0.39400 (0.43291)
	LR Test $\rho = \theta$	12.103	0.11233
	$p > \chi^2(df = 1)$	0.00503	0.73750
Attained Secondary	ρ	0.36132*** (0.09018)	0.49659*** (0.09164)
	θ	0.35623** (0.18284)	0.56249** (0.24232)
	$\rho - \theta$	0.00509 (0.21229)	0.06590 (0.24745)
	LR Test $\rho = \theta$	0.00317	0.01924
	$p > \chi^2(df = 1)$	0.95510	0.88967
Completed Primary	ρ	0.42096*** (0.11380)	0.65492*** (0.12240)
	θ	0.36004** (0.17838)	0.40604** (0.16784)
	$\rho - \theta$	0.06092 (0.24721)	0.24888 (0.20689)
	LR Test $\rho = \theta$	3.4608	1.4416
	$p > \chi^2(df = 1)$	0.06284	0.22990
Attained Primary	ρ	0.25271** (0.11211)	0.91825*** (0.31298)
	θ	0.52675*** (0.14222)	0.44113*** (0.14041)
	$\rho - \theta$	−0.27403 (0.21770)	0.47712 (0.31447)
	LR Test $\rho = \theta$	4.4609	2.9183
	$p > \chi^2(df = 1)$	0.03468	0.08758
No. Obs.		294	245

Notes: Standard errors, in parentheses, were recovered using standard approximation methods. White's heteroskedasticity correction was used. ***, **, * Significantly different from 0 at the 1% , 5%, and 10% levels, respectively.

Table A5: CES–Nested–in–CD Nonlinear Estimation (weighted–labor data)

Skilled Labor	Parameter	NLLS	NLLS with FE
Attained	θ	0.45075 (0.28610)	0.58865 (1.3369)
College	γ	0.72299*** (0.01680)	0.43746*** (0.03825)
Completed	θ	1.10118*** (0.28571)	0.94627*** (0.35455)
Secondary	γ	0.78039*** (0.01530)	0.59882*** (0.03739)
Attained	θ	1.11141*** (0.26531)	0.53203** (0.22830)
Secondary	γ	0.82371*** (0.01530)	0.71193*** (0.03336)
Completed	θ	0.89877*** (0.21526)	0.49696*** (0.17400)
Primary	γ	0.87246*** (0.01637)	0.81666*** (0.02749)
Attained	θ	0.71738*** (0.12867)	0.51553*** (0.13249)
Primary	γ	0.97742*** (0.06066)	0.99985*** (0.00246)
No. Obs.		294	245

Notes: Standard errors, in parentheses, were recovered using standard approximation methods. White's heteroskedasticity correction was used. ***, **, * Significantly different from 0 at the 1% , 5%, and 10% levels, respectively.

References

Arellano, Manuel, and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58 (April 1991), 277–297.

Barro, Robert J., and Jong-Wha Lee, "International Data on Educational Attainment: Updates and Implications," Oxford Economic Papers 53 (July 2001), 541–563.

Bergström, Villy, and Epaminondas E. Panas, "How Robust is the Capital–Skill Complementarity Hypothesis?," *Review of Economics and Statistics* 74 (August 1992), 540–546.

Berman, Eli, John Bound, and Stephen Machin, "Implications of Skill–Biased Technological Change: International Evidence," *Quarterly Journal of Economics* 113 (November 1998), 1245–1279.

Bils, Mark, and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review* 90 (December 2000), 1160–1183.

Blundell, Richard, and Stephen Bond, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87 (November 1998), 115–143.

Blundell, Richard, and Stephen Bond, "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews* 19:3 (2000), 321–340.

Caselli, Francesco, and Wilbur John Coleman II, "The World Technology Frontier," Working Paper, Harvard University (May 2002a).

Caselli, Francesco, and Wilbur John Coleman II, "The U.S. Technology Frontier," *American Economic Review* 92 (May 2002b), 148–152.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort, "Reopening the Convergence Debate: A New Look at Cross–Country Growth Empirics," *Journal of Economic Growth* 1 (September 1996), 363–389.

Duffy, John, and Chris Papageorgiou, "A Cross–Country Empirical Investigation of the Aggregate Production Function Specification," *Journal of Economic Growth* 5 (March 2000), 87–120.

Fallon, P.R., and Richard Layard, "Capital–Skill Complementarity, Income Distribution, and Output Accounting," *Journal of Political Economy* 83 (April 1975), 279–302.

Flug, Karnit, and Zvi Hercowitz, "Equipment Investment and the Relative Demand for Skilled Labor," *Review of Economic Dynamics* 3 (July 2000), 461–485.

Goldin, Claudia, and Lawrence F. Katz, "The Origins of Technology–Skill Complementarity," *Quarterly Journal of Economics* 113 (August 1998), 693–732.

Griliches, Zvi, "Capital–Skill Complementarity," *Review of Economics and Statistics* 51 (November 1969), 465–468.

Griliches, Zvi, and Jacques Mairesse, "Production Functions: The Search for Identification," in Steinar Strom, (Ed.) *Econometrics and Economic Theory in the 20th Century, The Ragnar Frisch Centennial Symposium*, Econometric Society Monograph No. 31, (Cambridge: Cambridge University Press 1998), 169–203.

Hamermesh, Daniel S., *Labor Demand*, 2nd Ed., (Princeton: Princeton University Press, 1993).

Islam, Nazrul, “Growth Empirics: A Panel Data Approach,” *Quarterly Journal of Economics* 110 (November 1995), 1127–1170.

Klenow, Peter J., and Andrés Rodríguez–Clare, “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” in Ben S. Bernanke and Julio J. Rotemberg, (Eds.), *NBER Macroeconomics Annual 1997*, (Cambridge, MA: MIT Press, 1997), 73–103.

Krusell, Per, Lee E. Ohanian, José–Victor Ríos–Rull, and Giovanni L. Violante, “Capital–Skill Complementarity and Inequality: A Macroeconomic Analysis,” *Econometrica* 68 (September 2000), 1029–1053.

Kumar, T. Krishna, and James H. Gapinski, “Nonlinear Estimation of the CES Production Parameters: A Monte Carlo Study,” *Review of Economics and Statistics* 56 (November 1974), 563–567.

Lee, Jong–Wha, “Length of Educational Attainment: A Cross–Country Data Set,” Unpublished, Korea University (2001).

Ruiz–Arranz, Marta, “Wage Inequality in the U.S.: Capital–Skill Complementarity Vs. Skill–Biased Technological Change,” Working Paper, Harvard University (November 2002).

Stokey, Nancy L., “Free Trade, Factor Returns, and Factor Accumulation,” *Journal of Economic Growth* 1 (December 1996), 421–447.

Sato, K., “A Two–Level Constant–Elasticity–of–Substitution Production Function,” *Review of Economic Studies* 34 (April 1967), 201–218.

Summers, Robert, and Alan Heston, “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988,” *Quarterly Journal of Economics* 106 (May 1991), 327–368.

Thursby, Jerry, “Alternative CES Estimation Techniques,” *Review of Economics and Statistics* 62 (May 1980), 295–299.