A Contribution to the Empirics of Economic Growth

Albert Alex Zevelev

May 6, 2011

1 Intoduction

This paper replicates Mankiw, Romer, and Weil's 1992 QJE paper "A Contribution to the Empirics of Economic Growth". Henceforth, we will refer to this paper as MRW. According to Google Scholar, MRW has 7328 citations and according to RePEc this is the 4th most cited paper in Economics.

First the paper tests the original (textbook) Solow growth model (SGM) (QJE, 1956), and shows that it fits the data pretty well. However the estimated coefficients on savings and labor force growth are larger then the model predicts.
Next the paper augments the SGM to include human capital, and we find that the estimated coefficients are very close to what the model predicts.

2 The Textbook Solow Model

The SGM seeks to explain economic growth taking the rates of saving "s", population growth "n" and technological progress "g" as exogenous (given outside the model). We assume a Cobb-Douglas (1928, AER) production function, so output (GNP) at time t is given by:

$$Y(t) = K(t)^{\alpha}(A(t)L(t))^{1-\alpha}, \text{ for } \alpha \in (0,1) \tag{2.1}$$

The notation is standard: Y is output, K capital, L labor, and A the level of technology. In this production function, technological change is labor augmenting which is known as "Harrod-neutral". (Note: since MRW, (Antras,

2004) has shown that the US production function is not Cobb-Douglas if we consider both capital augmenting and labor augmenting technological change. He estimates the elasticity of substitution to be less than 1.) L and A are assumed to grow exogenously at rates n and g:

$$L(t) = L(0)e^{nt} \tag{2.2}$$
$$A(t) = A(0)e^{gt} \tag{2.3}$$

The number of effective units of labor, $A(t)L(t) = A(0)L(0)e^{(n+g)t}$ grows at rate $n + g$. We will assume perfect competition, so the two inputs, capital and labor are paid their marginal products.
The model assumes that a constant fraction of output, s , is invested. The remaining output is consumed. We define k as the stock of capital per effective unit of labor,

$$k(t) \equiv \frac{K(t)}{A(t)L(t)}$$

and y as the level of output per effective unit of labor,

$$\begin{aligned} y(t) &= \frac{Y(t)}{A(t)L(t)} \\ &= K(t)^{\alpha}(A(t)L(t))^{-\alpha} \\ &= \left(\frac{K(t)}{(A(t)L(t))}\right)^{\alpha} \\ &= k(t)^{\alpha} \end{aligned}$$

The evolution of k is governed by:

$$\begin{aligned} \dot{k}(t) &= sy(t) - (n + g + \delta)k(t) \qquad (2.4) \\ &= sk(t)^{\alpha} - (n + g + \delta)k(t) \end{aligned}$$

where $\delta \in (0, 1)$ is the rate of capital depreciation. My intuition for this comes from the discrete version:

$$\underbrace{k(t+1)}_{\text{capital tomorrow}} = \underbrace{sk(t)^{\alpha}}_{\text{output saved}} + \underbrace{(1 - (n + g + \delta))k(t)}_{\text{output not depreciated}}$$

If we take the limit of $k(t+\Delta) - k(t)$ as $\Delta \to 0$, we get (2.4). The steady state is defined to be the level of capital k^{ss} at which $\dot{k}(t) = 0$, plugging this into equation (2.4) we have:

$$0 = s(k^{ss})^{\alpha} - (n+g+\delta)(k^{ss})$$

which implies

$$k^{ss} = \left(\frac{s}{n+g+\delta}\right)^{\frac{1}{1-\alpha}} \tag{2.5}$$

Now plugging in:

$$\begin{aligned}
y^{ss} &= (k^{ss})^{\alpha} \\
\frac{Y(t)}{A(t)L(t)} &= \left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}} \\
\frac{Y(t)}{L(t)} &= A(0)e^{gt}\left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}} \\
\log\left(\frac{Y(t)}{L(t)}\right) &= \log(A(0)) + gt + \left(\frac{\alpha}{1-\alpha}\right)\log(s) - \left(\frac{\alpha}{1-\alpha}\right)\log(n+g+\delta)
\end{aligned} \tag{2.6}$$

Equation (2.6) gives us an expression for steady-state income per-capita. Since capital's share in income α is roughly one third, the model implies elasticity of income per capita with respect to saving rate $\varepsilon_{\frac{Y}{L},s} = \frac{d\log(\frac{Y}{L})}{d\log(s)} = \frac{\alpha}{1-\alpha} \approx \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$, and an elasticity with respect to $(n+g+\delta)$ $\varepsilon_{\frac{Y}{L},s} = \frac{d\log(\frac{Y}{L})}{d\log(n+g+\delta)} = -\frac{\alpha}{1-\alpha} \approx -\frac{\frac{1}{3}}{\frac{2}{3}} = -\frac{1}{2}$.

3 Specification

The textbook Solow model predicts that real income is higher in countries with higher savings rates and lower in countries with higher values of $(n+g+\delta)$. We want to see if the data support this prediction. We assume that the growth rate of technology g and the rate of capital depreciation δ are constant across countries. In contrast the term $A(0)$ reflects not only the level of technology, but also resource endowments, climate, institutions and so on. It may differ across countries, so we assume that

$$\log(A(0)) = a + \varepsilon$$

where a is a constant and ε is a country specific shock. Thus log income per capita (at time 0) is:

$$\log\left(\frac{Y}{L}\right) = a + \left(\frac{\alpha}{1-\alpha}\right)\log(s) - \left(\frac{\alpha}{1-\alpha}\right)\log(n+g+\delta) + \varepsilon \qquad (3.1)$$

Equation (3.1) is our basic empirical specification in this section. We assume that s and n are independent of ε, so our regressors are exogenous. This implies we can estimate equation (3.1) with Ordinary Least Squares (OLS). NOTE: MRW point out that if s and n are endogenous and influenced by the level of income, then estimates of equation (3.1) by OLS are potentially inconsistent. To obtain consistent estimates, one needs to find instrumental variables that are correlated with s and n, but uncorrelated with the country specific shift in the production function ε. MRW say that finding such instrumental variables is a formidable task.
I personally believe that the savings rate can be influenced by the level of income, so I am concerned about endogeneity in this paper!

The data are from the Real National Accounts constructed by Summers and Heston [1988]. It includes real income, government and private consumption, investment, and population for almost all of the world except the centrally planned economies. The data are annual and cover 1960-1985. The population growth rate n is measured as the average rate of growth of the working age population, where working age is defined as 15-64. The savings rate s is measured as the average real investment (including government investment) in real GDP, and Y/L as real GDP in 1985 divided by the working age population in that year.
MRW considers three samples. The most comprehensive consists of all countries for which data are available other than those for which oil production is the dominant industry. This sample consists of 98 countries. MRW exclude the oil producers because the bulk of recorded GDP for these countries represents the extraction of existing resources, not value added; one should not expect standard growth models to account for measured GDP in these countries.

The second sample excludes countries whose population in 1960 was less than 1 million, or whose real income figures were based on extremely little primary data; measurement error is likely to be a greater problem for these countries. This sample consists of 75 countries.

The third sample consists of the 22 OECD countries with population greater than 1 million. This sample has high quality data and the variation in omitted country specific factors is small. A disadvantage is that this sample is small in size.

We estimate equation (3.1) with and without imposing the constraints that the coefficients on $\log(s)$ and $\log(n+g+\delta)$ are equal in magnitude and opposite in sign. We assume $g+\delta = 0.05$ because in the US $\delta \approx 0.03$ and $g \approx 0.02$.
We first estimate the unrestricted regression:

$$\log\left(\frac{Y}{L}\right) = \beta_0 + \beta_1 \log\left(\frac{I}{Y}\right) + \beta_2 \log(n+g+\delta) + \epsilon \qquad (3.2)$$

For the non-oil

```
#project for stat 521
rm(list=ls(all=TRUE))

library(foreign) #to read.dta
library(car)

cont <- read.dta("cont.dta")
attach(cont, pos=2)
x <- (log(iy/100) - log(workingagepop/100 + 0.05))
y <- (log(school/100) - log(workingagepop/100 + 0.05))
cont = cbind(cont,x,y)

contn <- subset(cont,(n==1))
conti <- subset(cont,(i==1))
conto <- subset(cont,(o==1))

## Non-oil, unrestricted regression
```

```
reg1 <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) , data=
print(summary(reg1))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05), data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
-1.79144 -0.39367  0.04124  0.43368  1.58046

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                     5.4299     1.5839   3.428 0.000900 ***
log(iy/100)                     1.4240     0.1431   9.951  < 2e-16 ***
log(workingagepop/100 + 0.05)  -1.9898     0.5634  -3.532 0.000639 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.6891 on 95 degrees of freedom
Multiple R-squared: 0.6009,Adjusted R-squared: 0.5925
F-statistic: 71.51 on 2 and 95 DF,  p-value: < 2.2e-16
```

For the Intermediate

```
## Intermediate, unrestricted regression
reg2 <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) , data=
print(summary(reg2))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05), data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-1.90004 -0.30528  0.05406  0.40391  1.36664

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept)                    5.3459     1.5431   3.464 0.000899 ***
log(iy/100)                    1.3176     0.1709   7.708 5.38e-11 ***
log(workingagepop/100 + 0.05)  -2.0172     0.5339  -3.778 0.000322 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.6106 on 72 degrees of freedom
Multiple R-squared: 0.5989,Adjusted R-squared: 0.5878
F-statistic: 53.76 on 2 and 72 DF,  p-value: 5.209e-15
```

For the OECD

```
## OECD, unrestricted regression
reg3 <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) , data=
print(summary(reg3))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05), data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.74350 -0.09797  0.09692  0.14802  0.58079

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    8.0206     2.5179   3.185  0.00487 **
log(iy/100)                    0.4999     0.4339   1.152  0.26357
log(workingagepop/100 + 0.05)  -0.7419     0.8522  -0.871  0.39484
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3774 on 19 degrees of freedom
Multiple R-squared: 0.1059,Adjusted R-squared: 0.01181
F-statistic: 1.126 on 2 and 19 DF,  p-value: 0.3452
```

Now we consider the restricted regression assuming $\beta_2 = -\beta_1$:

$$\log\left(\frac{Y}{L}\right) = \beta_0 + \beta_1\left(\log\left(\frac{I}{Y}\right) - \log(n+g+\delta)\right) + \epsilon \tag{3.3}$$

In our code we let $x \equiv (\log(s_k) - \log(n + g + \delta))$.
For the non-oil

```
## Non-oil, restricted regression
reg1b <- lm( log(gdpadult1985) ~ x , data=contn)
print(summary(reg1b))

Call:
lm(formula = log(gdpadult1985) ~ x, data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
-1.87388 -0.43133  0.03757  0.51698  1.49645

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.8724     0.1206   56.99   <2e-16 ***
x             1.4880     0.1247   11.93   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.6885 on 96 degrees of freedom
Multiple R-squared: 0.5974,Adjusted R-squared: 0.5932
F-statistic: 142.4 on 1 and 96 DF,  p-value: < 2.2e-16
```

For the Intermediate

```
## Intermediate, restricted regression
reg2b <- lm(log(gdpadult1985) ~ x , data=conti)
print(summary(reg2b))

Call:
lm(formula = log(gdpadult1985) ~ x, data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-2.01372 -0.30799 -0.01058  0.42530  1.29066

Coefficients:
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.0929     0.1456   48.71  < 2e-16 ***
x             1.4310     0.1391   10.29 7.58e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.6119 on 73 degrees of freedom
Multiple R-squared: 0.5917,Adjusted R-squared: 0.5861
F-statistic: 105.8 on 1 and 73 DF,  p-value: 7.576e-16
```

For the OECD

```
## OECD, restricted regression
reg3b <- lm(log(gdpadult1985) ~ x , data=conto)
print(summary(reg3b))

Call:
lm(formula = log(gdpadult1985) ~ x, data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.77381 -0.11080  0.09481  0.17503  0.57505

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.6244     0.5333  16.172 5.96e-13 ***
x             0.5538     0.3653   1.516    0.145
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3684 on 20 degrees of freedom
Multiple R-squared: 0.1031,Adjusted R-squared: 0.05824
F-statistic: 2.299 on 1 and 20 DF,  p-value: 0.1451
```

We wish to test the hypothesis $\beta_2 = -\beta_1$, with a linear hypothesis test. Let $R = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$ and $r = 0$, so
H_0: $R\beta = r$
H_1: $R\beta \neq r$
We test this hypothesis with the statistic:

$F = \frac{(SSR_R - SSR_U)/q}{SSR_U/(n-K)} \underset{H_0}{\sim} F_{q,n-K}$

For the non-oil (n = 98)

```
hypothesis.matrix <- matrix(c(0, 1, 1) , nrow=1 , ncol =3)

print(linearHypothesis( reg1, hypothesis.matrix, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     96 45.504
2     95 45.108  1   0.39612 0.8343 0.3634
```

For the Intermediate (n=75)

```
print(linearHypothesis( reg2, hypothesis.matrix, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     73 27.330
2     72 26.848  1   0.48226 1.2933 0.2592
```

For the OECD (n=22)

```
print(linearHypothesis( reg3, hypothesis.matrix, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) = 0

Model 1: restricted model
```

```
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     20 2.7147
2     19 2.7061  1 0.0085943 0.0603 0.8086
```

If the linear restriction holds then recall from equation (3.1) above that $\beta_1 = \frac{\alpha}{1-\alpha}$ which implies that $\alpha = \frac{\beta_1}{1+\beta_1}$.
For the non-oil $\alpha \approx 0.598$

For the Intermediate $\alpha \approx 0.5886$

For the OECD $\alpha \approx 0.3564$

Three aspects of the results support the Solow model. First, the coefficients on saving and population growth have the predicted signs, and for two of the three samples are highly significant. Second, the restriction that the coefficients on $\log(s)$ and $\log(n+g+\delta)$ are equal in magnitude and opposite in sign is not rejected in any of the samples. Third, and perhaps most important, differences in saving and population growth account for a large fraction of the cross-country variation in income per capita. For example, in the regression for the intermediate sample, the adjusted R^2 is 0.59. In contrast to the common claim that the Solow model "explains" cross-country variation in labor productivity largely by appealing to variations in technologies, the two readily observable variables on which the Solow model focuses in fact account for most of the variation in income per capita.

Nonetheless, the model is not completely successful. In particular, the estimated impacts of saving and labor force growth are much larger than the model predicts. The value of α implied by the coefficients should equal capital's share in income, which is roughly one third. The estimates, however, imply an α that is much higher. For example, the α implied by the coefficient in the constrained regression for the intermediate sample is 0.59 (with a standard error of 0.02). Thus, the data strongly contradict the prediction that $\alpha = 1/3$.

4 Adding Human Capital Accumulation to the Solow Model

We explore the effect of adding human capital accumulation to the Solow model. In the above section human capital is an omitted variable. We explore how adding it changes the empirical analysis in this section.
Let the production function be:

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} (A(t)L(t))^{1-\alpha-\beta} \tag{4.1}$$

where H denotes the stock of human capital and all other variables are defined as before. Let s_k be the fraction of income invested in physical capital and s_h be the fraction invested in human capital. The evolution of the economy is determined by:

$$\dot{k}(t) = s_k y(t) - (n + g + \delta)k(t) \tag{4.2}$$

$$\dot{h}(t) = s_h y(t) - (n + g + \delta)h(t) \tag{4.3}$$

where $y = Y/AL$, $k = K/AL$, and $h = H/AL$ are quantities per effective unit of labor. We are assuming one unit of consumption can be transformed costlessly into either one unit of physical capital or one unit of human capital. We are also assuming that human capital depreciates at the same rate as physical capital.
Note:

$$\begin{aligned} y &= \frac{Y}{AL} \\ &= \frac{K^{\alpha} H^{\beta} (AL)^{1-\alpha-\beta}}{AL} \\ &= K^{\alpha} H^{\beta} (AL)^{-\alpha-\beta} \\ &= K^{\alpha} H^{\beta} (AL)^{-\alpha} (AL)^{-\beta} \\ &= \left(\frac{K}{AL}\right)^{\alpha} \left(\frac{H}{AL}\right)^{\beta} \\ &= k^{\alpha} h^{\beta} \end{aligned}$$

We assume that $\alpha + \beta < 1$, which implies that there are decreasing returns to all capital. The steady state values of physical capital and human capital are k^{ss} and h^{ss} such that $\dot{k} = 0$ and $\dot{h} = 0$. So

$$\begin{aligned} s_k y(t) &= (n+g+\delta)k(t) \\ s_h y(t) &= (n+g+\delta)h(t) \end{aligned}$$

Or

$$\begin{aligned} k^{\alpha-1}h^{\beta} &= \frac{n+g+\delta}{s_k} \\ k^{\alpha}h^{\beta-1} &= \frac{n+g+\delta}{s_h} \end{aligned}$$

Now solving the top equation for h and plugging it into the second equation we get the steady state level of physical capital:

$$k^{ss} = \left(\frac{s_k^{1-\beta}s_h^{\beta}}{n+g+\delta}\right)^{\frac{1}{1-\alpha-\beta}} \tag{4.4}$$

Plugging this into the first equation, we get the steady state level of human capital:

$$h^{ss} = \left(\frac{s_k^{\alpha}s_h^{1-\alpha}}{n+g+\delta}\right)^{\frac{1}{1-\alpha-\beta}} \tag{4.5}$$

Plugging the steady state into the production function

$$\begin{aligned} y^* &= (k^*)^{\alpha}(h^*)^{\beta} \\ &= \left(\frac{s_k^{1-\beta}s_h^{\beta}}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha-\beta}} \left(\frac{s_k^{\alpha}s_h^{1-\alpha}}{n+g+\delta}\right)^{\frac{\beta}{1-\alpha-\beta}} \\ &= \left(\frac{s_k^{\alpha}s_h^{\beta}}{(n+g+\delta)^{\alpha+\beta}}\right)^{\frac{1}{1-\alpha-\beta}} \\ \frac{Y(t)}{A(t)L(t)} &= \left(\frac{s_k^{\alpha}s_h^{\beta}}{(n+g+\delta)^{\alpha+\beta}}\right)^{\frac{1}{1-\alpha-\beta}} \\ \frac{Y(t)}{L(t)} &= A(0)e^{gt}\left(\frac{s_k^{\alpha}s_h^{\beta}}{(n+g+\delta)^{\alpha+\beta}}\right)^{\frac{1}{1-\alpha-\beta}} \end{aligned}$$

Taking logs we get:

$$\log\left(\frac{Y(t)}{L(t)}\right) = \log(A(0)) + gt - \left(\frac{\alpha+\beta}{1-\alpha-\beta}\right)\log(n+g+\delta) \qquad (4.6)$$
$$+\left(\frac{\alpha}{1-\alpha-\beta}\right)\log(s_k) + \left(\frac{\beta}{1-\alpha-\beta}\right)\log(s_h)$$

This equation shows how income per capita depends on population growth and accumulation of physical and human capital. Similar to the textbook Solow model, the augmented model predicts coefficients in equation (4.6) that are functions of the factor shares. As before, α is physical capital's share of income, so we expect α to be about one third. Gauging a reasonable value of β is more difficult. In the US, the minimum wage – roughly the return to labor without human capital – has averaged 30 to 50 percent of the average wage in manufacturing. This suggests that 50 to 70 percent of total labor income represents the return to human capital, or that $\beta \in \left(\frac{1}{3}, \frac{1}{2}\right)$.

Note: my intuition for this is that since physical capital's share of income is one third, labor's share of income is two thirds. They argue that human capital represents 50 to 70 percent of total labor income, or $\frac{1}{2}\cdot\frac{2}{3} = \frac{1}{3}$ and $\frac{7}{10}\cdot\frac{2}{3} = \frac{7}{15}$. Consequently $\beta \in \left(\frac{1}{3}, \frac{1}{2}\right)$.

Equation (4.6) makes two predictions about the equations run in the textbook model in which human capital was ignored. First even if $\log(s_h)$ is independent of the other right-hand side variables, the coefficient on $\log(s_k)$ is greater than $\frac{\alpha}{1-\alpha}$. Because higher saving leads to higher income, it leads to a higher steady-state level of human capital, even if the percentage of income devoted to human-capital accumulation is unchanged. Hence, the presence of human-capital accumulation increases the impact of physical-capital accumulation on income.

Second, the coefficient on $\log(n+g+\delta)$ is larger in absolute value then the coefficient on $\log(s_k)$. In this model high population growth lowers income per capita because the amounts of both physical and human capital must be spread more thinly over the population.

An alternative way to express the role of human capital is by combining (4.6) with (4.5) yielding an equation for income as a function of the rate of investment in physical capital, the rate of population growth, and the *level* of human capital:

$$\log\left(\frac{Y(t)}{L(t)}\right) = \log(A(0)) + gt + \left(\frac{\alpha}{1-\alpha}\right)\log(s_k) - \left(\frac{\alpha}{1-\alpha}\right)\log(n+g+\delta) + \left(\frac{\beta}{1-\alpha}\right)\log(h^*) \tag{4.7}$$

Equation (4.7) is almost identical to the textbook Solow model where the level of human capital is a component of the error term. Because the saving and population growth rates influence h^*, one should expect human capital to be positively correlated with the saving rate and negatively correlated with population growth. Therefore, omitting the human-capital term biases the coefficients on saving and population growth.

MRW use a proxy for the rate of human capital accumulation s_h that measures the percentage of the working age population that is in secondary school. They call this proxy SCHOOL. Assuming SCHOOL is proportional to s_h they use it to estimate equation (4.6); the factor of proportionality will affect only the constant term.

The following equation describes the unrestricted regression:

$$\log\left(\frac{Y}{L}\right) = \beta_0 + \beta_1\log(s_k) + \beta_2\log(n+g+\delta) + \beta_3\log(School) + \varepsilon$$

For the non-oil

```
###### Solow Model with Human Capital
## Non-oil, unrestricted regression
reg1c <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) + log(
print(summary(reg1c))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05) + log(school/100), data = contn)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-1.2875 -0.3208  0.0726  0.3321  1.0952

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    6.84441    1.17745   5.813 8.37e-08 ***
log(iy/100)                    0.69671    0.13283   5.245 9.61e-07 ***
log(workingagepop/100 + 0.05) -1.74525    0.41594  -4.196 6.16e-05 ***
log(school/100)                0.65446    0.07271   9.001 2.44e-14 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.5077 on 94 degrees of freedom
Multiple R-squared: 0.7856,Adjusted R-squared: 0.7788
F-statistic: 114.8 on 3 and 94 DF,  p-value: < 2.2e-16
```

For the Intermediate

```
## Intermediate, unrestricted regression
reg2c <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) + log(
print(summary(reg2c))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05) + log(school/100), data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-1.12933 -0.30688  0.05654  0.24176  0.97441

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    7.79131    1.19242   6.534 8.30e-09 ***
log(iy/100)                    0.70037    0.15058   4.651 1.49e-05 ***
log(workingagepop/100 + 0.05) -1.49978    0.40322  -3.720 0.000396 ***
log(school/100)                0.73055    0.09523   7.671 6.79e-11 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.4547 on 71 degrees of freedom
```

```
Multiple R-squared: 0.7807,Adjusted R-squared: 0.7714
F-statistic: 84.25 on 3 and 71 DF,  p-value: < 2.2e-16
```

For the OECD

```
## OECD, unrestricted regression
reg3c <- lm(log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) + log(
print(summary(reg3c))

Call:
lm(formula = log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 +
    0.05) + log(school/100), data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.59596 -0.09879  0.03057  0.16791  0.63969

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                     8.6369     2.2143   3.901  0.00105 **
log(iy/100)                     0.2761     0.3889   0.710  0.48681
log(workingagepop/100 + 0.05)  -1.0755     0.7560  -1.423  0.17195
log(school/100)                 0.7676     0.2933   2.617  0.01746 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.33 on 18 degrees of freedom
Multiple R-squared: 0.3524,Adjusted R-squared: 0.2444
F-statistic: 3.264 on 3 and 18 DF,  p-value: 0.04552
```

In the restricted regression we estimate:

$$\log\left(\frac{Y}{L}\right) = \gamma_0 + \gamma_1 \left(\log(s_k) - \log(n+g+\delta)\right) + \gamma_2 \left(\log(School) - \log(n+g+\delta)\right) + \varepsilon$$

In our code we let $x \equiv (\log(s_k) - \log(n+g+\delta))$ and $y \equiv (\log(School) - \log(n+g+\delta))$. For the non-oil

```
## Non-oil, restricted regression
```

```
reg1d <- lm( log(gdpadult1985) ~ x + y , data=contn)
print(summary(reg1d))

Call:
lm(formula = log(gdpadult1985) ~ x + y, data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
-1.26287 -0.35150  0.06389  0.31553  1.04428

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.85309    0.14003  56.082  < 2e-16 ***
x            0.73828    0.12362   5.972 4.04e-08 ***
y            0.65708    0.07255   9.057 1.71e-14 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.507 on 95 degrees of freedom
Multiple R-squared: 0.7839,Adjusted R-squared: 0.7794
F-statistic: 172.3 on 2 and 95 DF,  p-value: < 2.2e-16
```

For the Intermediate

```
## Intermediate, restricted regression
reg2d <- lm(log(gdpadult1985) ~ x + y , data=conti)
print(summary(reg2d))

Call:
lm(formula = log(gdpadult1985) ~ x + y, data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-1.12830 -0.30568  0.04801  0.24058  0.96575

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.96624    0.15444  51.582  < 2e-16 ***
x            0.70908    0.13765   5.151 2.17e-06 ***
```

```
y            0.73304    0.09309   7.874 2.63e-11 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.4516 on 72 degrees of freedom
Multiple R-squared: 0.7806,Adjusted R-squared: 0.7745
F-statistic: 128.1 on 2 and 72 DF,  p-value: < 2.2e-16
```

For the OECD

```
## OECD, restricted regression
reg3d <- lm(log(gdpadult1985) ~ x + y , data=conto)
print(summary(reg3d))

Call:
lm(formula = log(gdpadult1985) ~ x + y, data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.59567 -0.10193  0.03436  0.16800  0.63997

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   8.7163     0.4662  18.697 1.08e-13 ***
x             0.2829     0.3339   0.847   0.4074
y             0.7686     0.2843   2.704   0.0141 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3212 on 19 degrees of freedom
Multiple R-squared: 0.3523,Adjusted R-squared: 0.2841
F-statistic: 5.167 on 2 and 19 DF,  p-value: 0.01614
```

Observe that the linear restriction is that $\beta_2 = -(\beta_1 + \beta_3)$. We test the hypothesis $\left[\begin{array}{cccc} 0 & 1 & 1 & 1 \end{array}\right] \beta = 0$:
For the non-oil

```
hypothesis.matrix2 <- matrix(c(0, 1, 1,1) , nrow=1 , ncol =4)
```

```
print(linearHypothesis( reg1c, hypothesis.matrix2, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) +
    log(school/100)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     95 24.418
2     94 24.226  1   0.19185 0.7444 0.3904
```

For the Intermediate

```
print(linearHypothesis( reg2c, hypothesis.matrix2, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) +
    log(school/100)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     72 14.684
2     71 14.680  1 0.0045263 0.0219 0.8828
```

For the OECD

```
print(linearHypothesis( reg3c, hypothesis.matrix2, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) ~ log(iy/100) + log(workingagepop/100 + 0.05) +
    log(school/100)
```

```
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     19 1.9604
2     18 1.9603  1 0.0001471 0.0014  0.971
```

We may find the implied α and β by solving the system:

$$\gamma_1 = \frac{\alpha}{1-\alpha-\beta}$$
$$\gamma_2 = \frac{\beta}{1-\alpha-\beta}$$

We obtain

$$\alpha = \frac{\gamma_1}{1+\gamma_1+\gamma_2}$$
$$\beta = \frac{\gamma_2}{1+\gamma_1+\gamma_2}$$

For the non-oil $\alpha \approx 0.31$ and $\beta \approx 0.28$

For the Intermediate $\alpha \approx 0.29$ and $\beta \approx 0.30$

For the OECD $\alpha \approx 0.14$ and $\beta \approx 0.37$

In the analysis above we see that the human capital measure enters significantly in all three samples. Including human capital greatly reduces the size of the coefficient on physical capital investment and improves the fit of the regression. These three variables explain almost 80 percent of the cross-country variation in income per capita in the non-oil and intermediate samples. Thus adding human capital to the Solow model improves its performance. The new parameters seem reasonable. Even using an imprecise proxy for human capital disposes a fairly large part of the model's residual variance.

5 Endogenous Growth and Convergence

Endogenous growth models are characterized by the assumption of nondecreasing returns to the set of reproducible factors of production. This assumption implies that countries that save more grow faster indefinitely and that countries need not converge in income per capita, even if they have the same preferences and technology.

Advocates of endogenous growth models attack an alleged empirical failure of the Solow model to explain cross-country differences: Barro [1989] says evidence indicates that per capita growth rates are uncorrelated with the starting level of per capita product.
We will reexamine this evidence on convergence to assess whether it contradicts the Solow model. We will do this by examining the predictions of the augmented Solow model for behavior out of the steady state.

The Solow model predicts that countries reach different steady states, as a function of accumulation of human and physical capital and population growth. Thus the Solow model does *not* predict convergence; it predicts only that income per capita in a given country converges to that country's steady state value. Thus the Solow model predicts convergence only after controlling for the determinants of the steady state, a phenomenon that might be called "conditional convergence".

In addition, the Solow model makes quantitative predictions about the speed of convergence to steady state. Let y^* be the steady-state level of income per effective worker given by equation (4.6), and let $y(t)$ be the actual value at time t. Approximating around the steady state, the speed of convergence is given by:

$$\frac{d\log(y(t))}{dt} = \lambda[\log(y^*) - \log(y(t))] \tag{5.1}$$

Where $\lambda = (n + g + \delta)(1 - \alpha - \beta)$
The model suggests a natural regression to study the rate of convergence. Equation (5.1) implies:

$$\log(y(t)) = (1 - e^{-\lambda t})\log(y^*) + e^{-\lambda t}\log(y(0)) \tag{5.2}$$

where $y(0)$ is income per effective worker at some initial date. Subtracting $\log(y(0))$ from both sides

$$\log(y(t)) - \log(y(0)) = (1 - e^{-\lambda t})\log(y^*) - (1 - e^{-\lambda t})\log(y(0)) \tag{5.3}$$

Substituting for y^*:

$$\begin{aligned}\log(y(t)) - \log(y(0)) = (1 - e^{-\lambda t})&\frac{\alpha}{1-\alpha-\beta}\log(s_k) \\ + (1 - e^{-\lambda t})&\frac{\beta}{1-\alpha-\beta}\log(s_h) \\ - (1 - e^{-\lambda t})&\frac{\alpha+\beta}{1-\alpha-\beta}\log(n+g+\delta) - (1 - e^{-\lambda t})\log(y(0))\end{aligned} \tag{5.4}$$

Thus in the Solow model the growth of income is a function of the determinants of the ultimate steady state and the initial level of income.

In endogenous growth models there is no steady state level of income; differences among countries in income per capita can persist indefinitely, even if the countries have the same saving and population growth rates.

The difference between regressions based on equation (5.4) and those presented earlier is that the latter regressions assume that a country has reached the steady state, or that deviations from the steady state are random.

First we study unconditional convergence with the following regression:

$$\log(Y_{1985}) - \log(Y_{1960}) = \beta_0 + \beta_1 \log(Y_{1960}) + \varepsilon \tag{5.5}$$

For the non-oil

```
## Non-oil
reg4 <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) , data=contn)
print(summary(reg4))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960),
    data = contn)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.09784 -0.27467 -0.02826  0.25975  1.17747 

Coefficients:
```

```
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -0.26658    0.37960  -0.702   0.4842
log(gdpadult1960)  0.09431    0.04962   1.901   0.0603 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.4405 on 96 degrees of freedom
Multiple R-squared: 0.03627,Adjusted R-squared: 0.02623
F-statistic: 3.613 on 1 and 96 DF,  p-value: 0.06033
```

For the Intermediate

```
## Intermediate
reg5 <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) , data=conti)
print(summary(reg5))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960),
    data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-1.04074 -0.31582  0.01031  0.24159  1.10531

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)        0.587524   0.432926   1.357    0.179
log(gdpadult1960) -0.004235   0.054835  -0.077    0.939

Residual standard error: 0.4077 on 73 degrees of freedom
Multiple R-squared: 8.17e-05,Adjusted R-squared: -0.01362
F-statistic: 0.005965 on 1 and 73 DF,  p-value: 0.9386
```

For the OECD

```
## OECD
reg6 <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) , data=conto)
print(summary(reg6))
```

```
Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960),
    data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.37984 -0.10050  0.03317  0.06571  0.47717

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        3.68629    0.68490   5.382 2.87e-05 ***
log(gdpadult1960) -0.34110    0.07852  -4.344 0.000315 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.183 on 20 degrees of freedom
Multiple R-squared: 0.4855,Adjusted R-squared: 0.4597
F-statistic: 18.87 on 1 and 20 DF,  p-value: 0.0003149
```

These regressions reproduce the results of many previous authors on the failure of incomes to converge [De Long 1988; Romer 1987]. The coefficient on the the initial level of income per capita is slightly positive for the non-oil sample and zero for the intermediate sample, and for both regressions the adjusted R^2 is essentially zero. There is no tendency for poor countries to grow faster on average than rich countries. We also see that there is a significant tendency toward convergence in the OECD sample. The coefficient on the initial level of income per capita is significantly negative, and the adjusted R^2 of the regression is 0.46.

We now study conditional convergence by adding the measure of investment and population growth to the RHS of the regression:

$$\log(Y_{1985}) - \log(Y_{1960}) = \beta_0 + \beta_1 \log(Y_{1960}) + \beta_2 \log(I/Y) + \beta_3 \log(n + g + \delta) + \varepsilon \tag{5.6}$$

For the non-oil

```
## Non-oil
reg4b <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) +log(iy/100)
```

```
print(summary(reg4b))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05), data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
-1.07648 -0.15215  0.01185  0.19595  0.96056

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    1.91938    0.83367   2.302  0.02352 *
log(gdpadult1960)             -0.14090    0.05202  -2.709  0.00803 **
log(iy/100)                    0.64724    0.08670   7.465 4.16e-11 ***
log(workingagepop/100 + 0.05) -0.30235    0.30438  -0.993  0.32311
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3507 on 94 degrees of freedom
Multiple R-squared: 0.4019,Adjusted R-squared: 0.3828
F-statistic: 21.05 on 3 and 94 DF,  p-value: 1.622e-10
```

For the Intermediate

```
## Intermediate
reg5b <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960)+ log(iy/100)
print(summary(reg5b))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05), data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-1.17569 -0.17325  0.00007  0.14674  0.93470

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept)                    2.24968    0.85472   2.632 0.010405 *
log(gdpadult1960)             -0.22783    0.05725  -3.980 0.000165 ***
log(iy/100)                    0.64587    0.10392   6.215 3.11e-08 ***
log(workingagepop/100 + 0.05) -0.45746    0.30743  -1.488 0.141176
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3258 on 71 degrees of freedom
Multiple R-squared: 0.3788,Adjusted R-squared: 0.3526
F-statistic: 14.43 on 3 and 71 DF,  p-value: 1.950e-07
```

For the OECD

```
## OECD
reg6b <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) + log(iy/100)
print(summary(reg6b))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05), data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.18289 -0.11598 -0.03232  0.06111  0.36293

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    2.14036    1.18068   1.813   0.0866 .
log(gdpadult1960)             -0.34991    0.06574  -5.322 4.65e-05 ***
log(iy/100)                    0.39010    0.17612   2.215   0.0399 *
log(workingagepop/100 + 0.05) -0.76624    0.34523  -2.220   0.0395 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.1529 on 18 degrees of freedom
Multiple R-squared: 0.6767,Adjusted R-squared: 0.6228
F-statistic: 12.56 on 3 and 18 DF,  p-value: 0.0001145
```

In all three samples the coefficient on the initial level of income is now significantly negative; there is strong evidence of conditional convergence. Moreover, the inclusion of investment and population growth rates improves substantially the fit of the regression.

We also add human capital to the regression:

$$\log(Y_{1985}) - \log(Y_{1960}) = \beta_0 + \beta_1 \log(Y_{1960}) + \beta_2 \log(I/Y) + \beta_3 \log(n+g+\delta) + \beta_4 \log(School) + \varepsilon \quad (5.7)$$

For the non-oil

```
## Non-oil
reg4c <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) +log(iy/100)
print(summary(reg4c))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05) + log(school/100),
    data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
-0.91041 -0.17599  0.01789  0.18439  0.93846

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    3.02152    0.82748   3.651 0.000431 ***
log(gdpadult1960)             -0.28837    0.06158  -4.683 9.62e-06 ***
log(iy/100)                    0.52374    0.08687   6.029 3.30e-08 ***
log(workingagepop/100 + 0.05) -0.50566    0.28861  -1.752 0.083061 .
log(school/100)                0.23112    0.05946   3.887 0.000190 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.327 on 93 degrees of freedom
Multiple R-squared: 0.4855,Adjusted R-squared: 0.4633
F-statistic: 21.94 on 4 and 93 DF,  p-value: 8.987e-13
```

For the Intermediate

```
## Intermediate
reg5c <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960)+ log(iy/100)
print(summary(reg5c))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05) + log(school/100),
    data = conti)

Residuals:
     Min       1Q   Median       3Q      Max
-0.97327 -0.17189  0.01514  0.16513  0.90109

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    3.70900    0.90878   4.081 0.000117 ***
log(gdpadult1960)             -0.36599    0.06743  -5.427 7.76e-07 ***
log(iy/100)                    0.53756    0.10229   5.255 1.52e-06 ***
log(workingagepop/100 + 0.05) -0.54499    0.28843  -1.890 0.062966 .
log(school/100)                0.27045    0.08037   3.365 0.001245 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3044 on 70 degrees of freedom
Multiple R-squared: 0.4653,Adjusted R-squared: 0.4348
F-statistic: 15.23 on 4 and 70 DF,  p-value: 5.262e-09
```

For the OECD

```
## OECD
reg6c <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) + log(iy/100)
print(summary(reg6c))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    log(iy/100) + log(workingagepop/100 + 0.05) + log(school/100),
    data = conto)

Residuals:
```

```
     Min       1Q   Median       3Q      Max
-0.21383 -0.07153 -0.02290  0.05779  0.31265

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                    2.75536    1.20102   2.294   0.0348 *
log(gdpadult1960)             -0.39769    0.07016  -5.668 2.78e-05 ***
log(iy/100)                    0.33179    0.17338   1.914   0.0727 .
log(workingagepop/100 + 0.05) -0.86341    0.33770  -2.557   0.0204 *
log(school/100)                0.22770    0.14501   1.570   0.1348
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.147 on 17 degrees of freedom
Multiple R-squared: 0.7176,Adjusted R-squared: 0.6512
F-statistic:  10.8 on 4 and 17 DF,  p-value: 0.0001525
```

We see that adding human capital further lowers the coefficient on the initial level of income, and it again improves the fit of the regression.

Finally we repeat the above estimation imposing the restriction that the coefficients on $\log(s_k)$, $\log(s_h)$ and $\log(n+g+\delta)$ sum to zero $\beta_2+\beta_3+\beta_4 = 0$.

$$\log(Y_{1985}) - \log(Y_{1960}) = \gamma_0 + \gamma_1 \log(Y_{1960}) + \gamma_2(\log(I/Y) - \log(n+g+\delta)) + \gamma_3(\log(School) - \log(n+g+\delta)) + \varepsilon \tag{5.8}$$

For the non-oil

```
## Non-oil
reg4d <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) + x + y, data
print(summary(reg4d))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    x + y, data = contn)

Residuals:
     Min       1Q   Median       3Q      Max
```

```
-0.87826 -0.17680  0.01979  0.17511  0.97166

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        2.45691    0.47297   5.195 1.19e-06 ***
log(gdpadult1960) -0.29790    0.06041  -4.931 3.51e-06 ***
x                  0.50067    0.08219   6.092 2.43e-08 ***
y                  0.23519    0.05916   3.975 0.000138 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3265 on 94 degrees of freedom
Multiple R-squared: 0.4816,Adjusted R-squared: 0.4651
F-statistic: 29.11 on 3 and 94 DF,  p-value: 2.121e-13
```

For the Intermediate

```
## Intermediate
reg5d <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960)+ x +y , data=
print(summary(reg5d))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    x + y, data = conti)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9442 -0.1638  0.0060  0.1711  0.9345

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        3.09033    0.52974   5.834 1.48e-07 ***
log(gdpadult1960) -0.37236    0.06687  -5.569 4.30e-07 ***
x                  0.50635    0.09508   5.325 1.13e-06 ***
y                  0.26569    0.08000   3.321  0.00142 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.3038 on 71 degrees of freedom
```

```
Multiple R-squared: 0.4599,Adjusted R-squared: 0.4371
F-statistic: 20.16 on 3 and 71 DF,  p-value: 1.486e-09
```

For the OECD

```
## OECD
reg6d <- lm( log(gdpadult1985)-log(gdpadult1960) ~ log(gdpadult1960) + x +y , data
print(summary(reg6d))

Call:
lm(formula = log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) +
    x + y, data = conto)

Residuals:
     Min       1Q   Median       3Q      Max
-0.22981 -0.08391 -0.01111  0.06970  0.27689

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        3.55369    0.63363   5.608 2.54e-05 ***
log(gdpadult1960) -0.40221    0.06918  -5.814 1.65e-05 ***
x                  0.39532    0.15174   2.605   0.0179 *
y                  0.24125    0.14244   1.694   0.1076
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.1454 on 18 degrees of freedom
Multiple R-squared: 0.7074,Adjusted R-squared: 0.6586
F-statistic:  14.5 on 3 and 18 DF,  p-value: 4.757e-05
```

We test the hypothesis $\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \end{bmatrix} \beta = 0$:
For the non-oil

```
hypothesis.matrix3 <- matrix(c(0,0, 1, 1,1) , nrow=1 , ncol =5)

print(linearHypothesis( reg4c, hypothesis.matrix3, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0
```

```
Model 1: restricted model
Model 2: log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) + log(iy/100) +
    log(workingagepop/100 + 0.05) + log(school/100)

  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     94 10.0196
2     93  9.9455  1  0.074066 0.6926 0.4074
```

For the Intermediate

```
print(linearHypothesis( reg5c, hypothesis.matrix3, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) + log(iy/100) +
    log(workingagepop/100 + 0.05) + log(school/100)

  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     71 6.5526
2     70 6.4874  1  0.065198 0.7035 0.4045
```

For the OECD

```
print(linearHypothesis( reg6c, hypothesis.matrix3, rhs=0))

Hypothesis:
log(iy/100) + log(workingagepop/100  + 0.05) + log(school/100) = 0

Model 1: restricted model
Model 2: log(gdpadult1985) - log(gdpadult1960) ~ log(gdpadult1960) + log(iy/100) +
    log(workingagepop/100 + 0.05) + log(school/100)

  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     18 0.38077
2     17 0.36742  1  0.013345 0.6174 0.4428
```

We find that this restriction is not rejected and that imposing it has little effect on the coefficients.

In general we have seen that departures from steady state represent a larger share of cross-country variation in income per capita for the OECD than for the broader samples. If the OECD countries are far from their steady states, then population growth and capital accumulation have not yet had their full impact on standards of living; hence we obtain lower estimated coefficients and lower R^2's for the OECD in the specifications that do not consider out-of-steady-state dynamics. Similarly, the greater importance of departures from steady state for the OECD would explain the finding of greater unconditional convergence.

The bottom line: after controlling for those variables that the Solow model says determine the steady state, there is substantial convergence in income per capita. Moreover, convergence occurs at approximately the rate that the model predicts.

6 conclusion

This paper suggested that international differences in income per capita are best understood using an augmented Solow growth model. In this model output is produced from physical capital, human capital, and labor, and is used for investment in physical capital, investment in human capital and consumption. One production function that is consistent with these empirical results is $Y = K^{1/3}H^{1/3}L^{1/3}$.

This model of growth has several implications. First, the elasticity of income with respect to the stock of physical capital is not substantially different from capital's share in income. This indicates that capital receives approximately its social return. There are not substantial externalities to the accumulation of physical capital.

Second, despite the absence of externalities, the accumulation of physical capital has a larger impact on income per capita than the textbook Solow model implies. A higher saving rate leads to higher income in steady state, which in turn leads to a higher level of human capital, even if the rate of human-capital accumulation is unchanged.