Maurer, Rainer

Working Paper

OLS-Estimation of conditional and unconditional sigma- and beta-convergence of per capita income: Implications of Solow-Swan and Ramsey-Cass models

Kiel Working Papers, No. 698

Provided in cooperation with:
Institut für Weltwirtschaft (IfW)

Kieler Arbeitspapiere
Kiel Working Papers

Kiel Working Paper No. 698

OLS-Estimation of Conditional and Unconditional Sigma- and Beta-Convergence of Per Capita Income

Implications of Solow-Swan and Ramsey-Cass models

by Rainer Maurer

Institut für Weltwirtschaft an der Universität Kiel
The Kiel Institute of World Economics

ISSN 0342 - 0787

Kiel Institute for World Economics

Düsternbrooker Weg 120

D- 24105 Kiel

Kiel Working Paper No. 698

OLS-Estimation of Conditional and Unconditional Sigma- and Beta-Convergence of Per Capita Income

Implications of Solow-Swan and Ramsey-Cass models

by Rainer Maurer

June 1995

Abstract

In this paper I discuss the general statistical relationships between beta- and sigma-convergence (for a definition see section 2) and the implications of the Solow-Swan and Ramsey-Cass model for an OLS-estimation of beta- and sigma-convergence of the log of per capita GDP over a cross section of countries. Furthermore, I present tests of conditional and unconditional sigma- and beta-convergence.

The discussion of the statistical relations exhibits that based on the Cauchy-Schwarz inequality it is possible to show that sigma-convergence implies necessarily beta-convergence but that beta-convergence is compatible with sigma-convergence as well as sigma-divergence. The discussion of the implications of the Solow-Swan model shows that - depending on identical stochastics - these models imply *unconditional* beta- and sigma-convergence, if the cross section sample includes only economies with identical steady state parameters. If the economies display different steady state parameters both models imply *conditional* beta- and sigma-convergence.

A replication of the well-known test results for conditional beta-convergence based on the Summers/Heston (1991) and the Barro/Lee (1993) data sets, does not reject conditional beta-convergence. However, the results of the tests for conditional sigma-convergence are sensitive concerning slight modifications of the cross section sample of countries.

Key words: Beta- and sigma convergence of per capita GDP, Solow-Swan growth model, Ramsey growth model, multicollinearity, BLUE property of OLS-estimators, empirical test.

JEL-Classification: C12, C21, F43, O11,

Contents:

1. Introduction

The question whether the implications of the Solow-Swan model and the Ramsey-Cass model concerning conditional beta-convergence are supported by the data of a cross section of heterogeneous countries (such as the Summers/Heston (1991) data set) has become an issue.[1] Some authors like Quah (1993a), (1993b) argue that the results of Barro (1991), Barro/Sala-i-Martin (1991), (1992) and Mankiw/Romer/Weil (1992) tend to be a statistical artifact plagued by Galtons-regression-fallacy. Quah (1993a) shows that beta-convergence does not necessarily imply a reduction of the cross-section sample variance of per capita income (sigma-convergence) but is indeed compatible with a growing cross-section sample variance (sigma-divergence).

I complement the statistical relations between beta- and sigma-convergence. I show that the Cauchy-Schwarz inequality implies that although beta-convergence does not imply sigma-convergence, sigma-convergence does imply beta-convergence. Hence, empirical findings of sigma-convergence across the U.S. states, Japanese prefectures or European regions (Barro/Sala-i-Martin (1995), chapter 11) does imply the existence of beta-convergence .

A formal analysis of the problems of an OLS-estimation of conditional and unconditional beta-convergence does not indicate that such an estimation is „uninformative". I find that a major problem of a consistent estimation of conditional beta-convergence stems from a multicollinearity problem, which is caused by deficient information on the initial states of the economies. However, I provide arguments that in spite of these difficulties, the Solow-Swan model still implies that an OLS-estimation should yield beta-convergence. Hence, empirical evidence of beta-divergence would reject the Solow-Swan model. Therefore an OLS-test for beta-convergence is „informative". The econometric reason, why the multicollinearity problem can be overcome, is the fact that OLS-estimators keep their BLUE-property even in the presence of strong multicollinearity, if the other assumptions of the Gauß-Markov theorem hold. I argue that the variance inflating effect of multicollinearity is a minor problem, if the estimation yields results that are significant although standard errors are inflated.

However, as beta-convergence may go hand in hand with sigma-constancy or sigma-divergence, a test for sigma-convergence yields additional information. Therefore, I analyze the implications of the Solow-Swan model on sigma-convergence. I find that (depending on

---

1 As the implications of the Solow-Swan model and the Ramsey-Cass model on beta- and sigma-convergence are identical, I quote within the following only the Solow-Swan model.

the stochastics of the model) the model implies unconditional sigma-convergence, if the steady state parameters of the cross section of economies are similar, and that the model implies conditional sigma-convergence, if the steady state parameters of the cross section of economies are different.

The empirical results (based on the Summers/Heston (1991) data) replicate the well known finding that conditional beta-convergence is not rejected by the data, but show that the results on conditional sigma-convergence are sensitive to small alternations of the country sample.

This paper is organized as follows: Section 2 overviews the general statistical relations between sigma- and beta-convergence. Section 3 discusses the implications of the Solow-Swan model for an OLS-estimation of unconditional and conditional beta-convergence and replicates the empirical results of a test for conditional beta-convergence. Section 4 discusses the implications of the Solo-Swan model on conditional and unconditional sigma-convergence and presents the empirical results of a test for conditional sigma-convergence. Section 5 draws conclusions.

2. General statistical relations between sigma- and beta-convergence

In this section I discuss the general statistical relations between sigma- and beta-convergence. I show that given the very general assumptions of the Cauchy-Schwarz inequality, several relations must hold: Suppose a sample of N variables, $Y_{i,t}$ with $i \in N$, that move through time following an unknown law of motion or no law at all. Given a time series of the realizations of this sample, it is possible to test whether its law of motion displays some characteristics of motion or not.

It is for example possible to test whether the per period cross section variance of this sample, $\mathrm{var}(y_{i,t})$, grows or decreases over time, where $y_{i,t} = \ln(Y_{i,t})$.[2] Following a common definition, if $\mathrm{var}_i(y_{i,t})$ grows over time this is called sigma-divergence; if $\mathrm{var}_i(y_{i,t})$ decreases over time this is called sigma-convergence (Barro/Sala-i-Martin (1995), chap. 11).

[2] The development of the time series behaviour of $\mathrm{var}_i(Y_i)_t$ is not very informative, as it grows/decreases over time if Y_i grows/decreases: $Y_{i,t} = Y_{i,t=0}\, e^{\beta t} \Rightarrow \mathrm{var}_i(Y_i)_t = \mathrm{var}_i(Y_i)_{t=0}\, e^{2\beta t}$. To get information on the relative degree of convergence or divergence it is necessary to take logs or to normalize by the mean of Y_i. Therefore the discussion on beta- and sigma-convergence in the empirical literature on growth theory is based on the log of per capita GDP (see Barro/Sala-i-Martin (1995), chap. 11).

Another possibility to characterize the time series behavior of the sample, is to ask whether variables with a small value grow faster or slower than variables with a high value, i. e. whether the cross section covariance $\operatorname{cov}((y_{i,t_1} - y_{i,t_0}), y_{i,t_0})$ holds the following inequality: $\operatorname{cov}((y_{i,t_1} - y_{i,t_0}), y_{i,t_0}) < 0 \iff \operatorname{cov}(y_{i,t_1}, y_{i,t_0}) < \operatorname{var}(y_{i,t_0})$. Following a common definition, this is called beta-convergence (Barro/Sala-i-Martin (1995), chap. 11). Consequently, beta-divergence is defined as $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) > \operatorname{var}(y_{i,t_0})$.

As it turns out, based on these definitions six lemmas can be proven, given the Cauchy-Schwarz inequality[3] ($\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) \leq \sqrt{\operatorname{var}(y_{i,t_1})}\sqrt{\operatorname{var}(y_{i,t_0})}$):[4]

Lemma 1: Sigma-convergence implies necessarily beta-convergence.

Suppose sigma-convergence holds. This implies $\operatorname{var}(y_{i,t_1}) < \operatorname{var}(y_{i,t_0}) \Rightarrow \sqrt{\operatorname{var}(y_{i,t_1})}\sqrt{\operatorname{var}(y_{i,t_0})} < \sqrt{\operatorname{var}(y_{i,t_0})}\sqrt{\operatorname{var}(y_{i,t_0})}$. Inserting this in the Cauchy-Schwarz inequality yields $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) < \operatorname{var}(y_{i,t_0})$, which is the definition of beta-convergence.

Lemma 2: Beta-divergence implies necessarily sigma-divergence.

Suppose beta-divergence holds. This implies $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) > \operatorname{var}(y_{i,t_0})$. The Cauchy-Schwarz inequality implies $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) \leq \sqrt{\operatorname{var}(y_{i,t_1})}\sqrt{\operatorname{var}(y_{i,t_0})}$. Combining these two inequalities yields $\operatorname{var}(y_{i,t_1}) > \operatorname{var}(y_{i,t_0})$, which is the definition of sigma-divergence.

Lemma 3: Beta-convergence is compatible with sigma-convergence or sigma-divergence.

Suppose beta-convergence holds. This implies $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) < \operatorname{var}(y_{i,t_0})$. The Cauchy-Schwarz inequality implies $\operatorname{cov}(y_{i,t_1}, y_{i,t_0}) \leq \sqrt{\operatorname{var}(y_{i,t_1})}\sqrt{\operatorname{var}(y_{i,t_0})}$. Consequently $\operatorname{var}(y_{i,t_0}) < \sqrt{\operatorname{var}(y_{i,t_1})}\sqrt{\operatorname{var}(y_{i,t_0})} \Leftrightarrow \operatorname{var}(y_{i,t_0}) < \operatorname{var}(y_{i,t_1})$, i.e. sigma-divergence, as well

---

[3] The Cauchy-Schwarz inequality is the reason why the absolut value of a correlation coefficient is always smaller than unity. It holds for any two dimensional random variables, (y_{t_1}, y_{t_2}), over an normally behaved probability space. The proof can be found in any text book on statistics.

[4] Lemma 4 and 6 is shown in Quah (1993a), p. 7.

as $\mathrm{var}(y_{i,t_0}) > \sqrt{\mathrm{var}(y_{i,t_1})}\sqrt{\mathrm{var}(y_{i,t_0})} \Leftrightarrow \mathrm{var}(y_{i,t_0}) > \mathrm{var}(y_{i,t_1})$, i.e. sigma-convergence, may hold.

Lemma 4: Sigma-divergence is compatible with beta-divergence or beta-convergence.

Suppose sigma-divergence holds. This implies $\mathrm{var}(y_{i,t_1}) > \mathrm{var}(y_{i,t_0}) \Rightarrow \sqrt{\mathrm{var}(y_{i,t_1})}\sqrt{\mathrm{var}(y_{i,t_0})} > \mathrm{var}(y_{i,t_0})$. The Cauchy-Schwarz inequality implies $\mathrm{cov}(y_{i,t_1}, y_{i,t_0}) \le \sqrt{\mathrm{var}(y_{i,t_1})}\sqrt{\mathrm{var}(y_{i,t_0})}$. Consequently $\mathrm{var}(y_{i,t_0}) < \mathrm{cov}(y_{i,t_1}, y_{i,t_0})$, i.e. beta-convergence, as well as $\mathrm{var}(y_{i,t_0}) > \mathrm{cov}(y_{i,t_1}, y_{i,t_0})$, i.e. beta-divergence may hold.

Lemma 5: Beta-constancy is compatible with sigma-convergence or sigma-constancy

Suppose beta-constancy holds. This implies $\mathrm{var}(y_{i,t_0}) = \mathrm{cov}(y_{i,t_1}, y_{i,t_0})$. The Cauchy-Schwarz inequality implies $\mathrm{cov}(y_{i,t_1}, y_{i,t_0}) \le \sqrt{\mathrm{var}(y_{i,t_1})}\sqrt{\mathrm{var}(y_{i,t_0})}$. Combining both relations yields $\mathrm{var}(y_{i,t_0}) \le \mathrm{var}(y_{i,t_1})$, which is compatible with sigma-convergence as well as sigma-constancy.

Lemma 6: Sigma-constancy is compatible with beta-convergence or beta-constancy

Suppose sigma-convergence holds. This implies $\mathrm{var}(y_{i,t_0}) = \mathrm{var}(y_{i,t_1})$. This implies $\sqrt{\mathrm{var}(y_{i,t_0})}\,\sqrt{\mathrm{var}(y_{i,t_0})} = \sqrt{\mathrm{var}(y_{i,t_1})}\,\sqrt{\mathrm{var}(y_{i,t_0})}$. The Cauchy-Schwarz inequality implies $\mathrm{cov}(y_{i,t_1}, y_{i,t_0}) \le \sqrt{\mathrm{var}(y_{i,t_1})}\sqrt{\mathrm{var}(y_{i,t_0})}$. Combining both relations yields $\mathrm{cov}(y_{i,t_1}, y_{i,t_0}) \le \mathrm{var}(y_{i,t_0})$, which is compatible with beta-convergence as well as beta-constancy.

An empirical demonstration of lemma one is given by the results of Barro/Sala-i-Martin (1995), figure 11.2 and figure 11.4, on beta- and sigma convergence of per capita income across the U.S. states from 1890-1990: Figure 11.4 displays the existence of strong sigma-convergence of the per capita income of the U.S. This implies according to lemma 1 necessarily beta-convergence. Therefore the actual existence of beta-convergence of per capita income across the U.S. states, which is displayed by figure 11.2 of Barro/Sala-i-Martin (1995), is not surprising. In fact, given lemma 1, figure 11.2 contains no additional information compared to figure 11.4. However, as lemma 3 implies, the other direction holds not: Figure 11.2 does not imply figure 11.4. Consequently, figure 11.4 contains additional information compared to figure 11.2. Hence, both figures are dramaturgically perfectly sequenced.

3. Estimation of conditional and unconditional beta-convergence: Implications of the Solow-Swan model

In this section I discuss the implications of the Solow-Swan model for an OLS-estimation of beta-convergence. As shown in Barro/Sala-i-Martin (1995), p.27 and p. 80, the Solow-Swan model implies the following relation to hold:[5]

$$y_{i,t_1} = \left(1-e^{-\beta t_1}\right) y_i^* + e^{-\beta t_1} y_{i,t_0} \tag{1}$$

where y_{i,t_1} equals the log of per capita GDP of country i of period t_1, y_{i,t_0} equals the log of per capita GDP of the initial state of economy i, y_i^* equals the steady state level of the log of per capita GDP and β equals $(1-\alpha)(g+n+d)$, where α equals the production elasticity of accumulating capital, g equals the rate of technological progress, n equals the growth rate of labor supply and d the rate of capital depreciation.[6] Equation (1) is derived under the assumption of diminishing marginal returns of accumulating capital, such that $0 < \alpha < 1$. As $(g+n+d)$ has a positive value, equation (1) implies that the economy approaches its steady state level as $t \to \infty$ with a speed that is the higher the higher is β.

One important implication of equation (1) is the implication of beta-convergence. To see this subtract y_{i,t_0} from both sides to get:

$$y_{i,t_1} - y_{i,t_0} = \left(1-e^{-\beta t_1}\right) y_i^* - \left(1-e^{-\beta t_1}\right) y_{i,t_0} \tag{2}$$

As the negative sign implies, the lower y_{i,t_0} the higher the growth rate on the left hand side of the equation.[7] To test for this implication of the model a simple single variable OLS-regression of y_{i,t_1} on y_{i,t_0} or $\left(y_{i,t_1} - y_{i,t_0}\right)$ on y_{i,t_0} would yield a consistent estimate of $e^{-\beta t}$ resp. $\left(1-e^{-\beta t}\right)$, if and only if y_{i,t_0} were known. This can be shown based on a regression of equation (1) as well as equation (2), as an OLS-regression is indifferent to linear

---

5 Equation (1) can be derived by a Taylor expansion of the differential equations derived from the Solow/Swan or Ramsey model (see Barro/Sala-i-Martin (1995), equations (1.30) and (2.33)).

6 The per capita units of $y_{i,t}$ are measured in efficiency units, which implies $y_{i,t} = Y_{i,t}/A_t L_t$, where A_t is the measure of labour productivity.

7 Here and in the following I neclegt the division of the growth rate through the number of periods $(t_1\text{-}t_0)$, as this has only a scale effect on the magnitudes of coefficients and does not qualitatively influence the results.

transformations. Therefore, in the following I concentrate the analysis on equation (1). Hence, consider the following cross country regression of equation (1):

$$y_{i,t_1} = \alpha + \beta_1 \; y_{i,t_0} + \mu_i \qquad (3)$$

The OLS-estimator of β_1 equals then:

$$\beta_1 = \frac{\operatorname{cov}(y_{i,t_1}, y_{i,t_0})}{\operatorname{var}(y_{i,t_0})} \qquad (4)$$

Inserting equation (1) augmented by a normally behaved error term $\varepsilon_{i,t}$, with $\varepsilon_{i,t} \approx N(0,\sigma^2)$, $\operatorname{cov}(\varepsilon_{i,t}, \varepsilon_{i,t-n}) = \operatorname{cov}(\varepsilon_{i,t}; \varepsilon_{j,t}) = 0$, $\operatorname{cov}(\varepsilon_{i,t}, y_{i,t_0}) = \operatorname{cov}(\varepsilon_{i,t}, y^*) = 0$, yields:

$$\beta_1 = \frac{\left(1 - e^{-\beta t_1}\right)\operatorname{cov}(y_i^*, y_{i,t_0}) + e^{-\beta t_1}\operatorname{var}(y_{i,t_0}) + \operatorname{cov}(\varepsilon_{i,t}, y_{i,t_0})}{\operatorname{var}(y_{i,t_0})} \qquad (5)$$

As the model implies that the steady state value of per capita GDP, y_i^*, is independent from its initial value, y_{i,t_0}, and under the assumption that the country specific shock in period t is also independent from y_{i,t_0}, both covariances are zero, such that β_1 equals indeed $e^{-\beta t_1}$. It is important to note that this is also the case, if var (y_i^*) > 0, i.e. if the steady state parameters of the countries diverge. Hence, under these assumptions an OLS-estimation would indeed yield consistent estimates of $e^{-\beta t_1}$. As $0 < e^{-\beta t_1} < 1$ holds under the null hypothesis that the Solow-Swan model is true, this implies the following inequalities to hold

$$0 < \beta_1 = e^{-\beta t_1} = \frac{\operatorname{cov}(y_{i,t_1}, y_{i,t_0})}{\operatorname{var}(y_{i,t_0})} < 1 \Leftrightarrow \operatorname{cov}(y_{i,t_1}, y_{i,t_0}) < \operatorname{var}(y_{i,t_0}) \qquad (6)$$

$$\Rightarrow \operatorname{cov}(y_{i,t_1}, y_{i,t_0}) - \operatorname{var}(y_{i,t_0}) < 0 \qquad (7)$$

$$\Leftrightarrow \operatorname{cov}\left((y_{i,t_1} - y_{i,t_0}), y_{i,t_0}\right) < 0 \qquad (8)$$

Equation (8) is the definition for beta-convergence. Hence, this is just another way to show that the Solow-Swan model implies beta-convergence. However, the typical problem of an OLS-estimation of β_1 is that y_{i,t_0} is unknown. Therefore, it has to be approximated by the value of per capita GDP in some base period. Yet, as equation (1) implies this proxy for y_{i,t_0} is typically influenced by the steady state value of per capita GDP, y_i^*. Table 1 shows that this implication of the Solow-Swan model is also found in the data: The level of per capita GDP of „base period" 1960 is strongly correlated with variables that are typically used as proxies for the steady state parameters of the Solow-Swan model.

Table 1 Correlation matrix of per capita GDP and typical steady state parameters[1]

	y85	y60	n	s_k	h^*	s_h	ROA	RAIL	GOV	ASS	COU
y85	1,00										
y60	0,92	1,00									
n	-0,67	-0,64	1,00								
s_k	0,61	0,48	-0,32	1,00							
h^*	0,83	0,77	-0,57	0,60	1,00						
s_h	0,78	0,76	-0,61	0,47	0,77	1,00					
ROA	0,53	0,45	-0,54	0,42	0,57	0,49	1,00				
RAIL	0,49	0,39	-0,55	0,27	0,52	0,46	0,77	1,00			
GOV	-0,56	-0,52	0,30	-0,37	-0,50	-0,34	-0,42	-0,36	1,00		
ASS	-0,06	-0,01	0,08	-0,16	-0,06	-0,18	0,05	0,02	-0,11	1,00	
COU	-0,23	-0,21	0,19	-0,24	-0,08	-0,19	-0,16	-0,10	0,06	0,16	1,00

[1] The numbers are correlation coefficients. 'Y85' = log of per capita GDP 1985, 'Y60' = log of per capita GDP 1960, 'n' = population growth rate 1980-85, ' s_k ' = share of real investments in GDP 1985, ' s_h ' = percentage of „secondary school attained" in the total population in 1985, ' h^* ' = average schooling years in the total population over age 25 in 1985, 'ROAD' = length of road network per km^2 of country area in 1985, 'RAIL' = lenght of rail network per km^2 of country area in 1985, 'GOV' = share of government expenditures on GDP in 1985, 'ASS' = number of assassinations per million population per year in 1980-85, 'COU' = number of coups per year in 1980-85. Source: Summers/Heston (1991), Barro/Lee (1993) and IRTU (1990).

As it turns out, this can give raise to some estimation problems. In order to analytically derive the consequences of this problem, consider the following notation that replicates equation (1) for two different points in time:

$$y_{i,t_1} = a_1\, y_i^* + b_1\, y_{i,t_0} + \varepsilon_{i,t_1} \quad \text{with} \quad a_1 = 1 - e^{-\beta t_1} \text{ and } b_1 = e^{-\beta t_1} \qquad (9)$$

$$y_{i,t_2} = a_2\, y_i^* + b_2\, y_{i,t_0} + \varepsilon_{i,t_2} \quad \text{with} \quad a_2 = 1 - e^{-\beta t_2} \text{ and } b_2 = e^{-\beta t_2} \qquad (10)$$

As $t_1 < t_2$, the following relations must hold:

$$a_1 < a_2 \text{ and } b_1 > b_2 \qquad (11)$$

Based on equations (9) and (10) and on some basic statistical lemmas, the relations given in the following box 1 hold:

Box 1 - Some statistical relations following equations (9) and (10)

$$\operatorname{cov}\left(y_{i,t_1}, y_i^*\right) = a_1 \operatorname{var}\left(y_i^*\right)$$

$$\operatorname{cov}\left(y_{i,t_2}, y_i^*\right) = a_2 \operatorname{var}\left(y_i^*\right)$$

$$\operatorname{cov}\left(y_{i,t_1}, y_0\right) = b_1 \operatorname{var}\left(y_0\right)$$

$$\operatorname{cov}\left(y_{i,t_2}, y_0\right) = b_2 \operatorname{var}\left(y_0\right)$$

$$\operatorname{cov}\left(y_{i,t_1}, y_{i,t_2}\right) = a_1 a_2 \operatorname{var}\left(y_i^*\right) + b_1 b_2 \operatorname{var}\left(y_0\right)$$

$$\operatorname{cov}\left(y_{i,t_1}, y_{i,t_2}\right) = a_1 \operatorname{cov}\left(y_{i,t_2}, y_i^*\right) + b_1 \operatorname{cov}\left(y_{i,t_2}, y_0\right)$$

$$\operatorname{cov}\left(y_{i,t_1}, y_{i,t_2}\right) = a_2 \operatorname{cov}\left(y_{i,t_1}, y_i^*\right) + b_2 \operatorname{cov}\left(y_{i,t_1}, y_0\right)$$

$$\operatorname{var}\left(y_{i,t_1}\right) = a_1^2 \operatorname{var}\left(y_i^*\right) + b_1^2 \operatorname{var}\left(y_0\right) + \sigma^2$$

$$\operatorname{var}\left(y_{i,t_2}\right) = a_2^2 \operatorname{var}\left(y_i^*\right) + b_2^2 \operatorname{var}\left(y_0\right) + \sigma^2$$

Consider now an OLS-estimation of the following equation:

$$y_{i,t_2} = \alpha + \beta_1 \; y_{i,t_1} + \mu_i \tag{12}$$

This equation equals equation (3) with the exception that y_{i,t_0} is approximated by some y_{i,t_1}. The OLS-estimator of β_1 equals then:

$$\beta_1 = \frac{\operatorname{cov}\left(y_{i,t_2}, y_{i,t_1}\right)}{\operatorname{var}\left(y_{i,t_1}\right)} \tag{13}$$

Inserting equations (8) and (9) and using some of the relations of box 1 yields:

$$\beta_1 = \frac{a_1 a_2 \operatorname{var}\left(y_i^*\right) + b_1 b_2 \operatorname{var}\left(y_{i,t_0}\right)}{a_1^2 \operatorname{var}\left(y_i^*\right) + b_1^2 \operatorname{var}\left(y_{i,t_0}\right) + \sigma^2} \tag{14}$$

If all economies $i \in N$ had the same steady state parameters, the steady state level of per capita GDP would be equal in all countries, such that $\operatorname{var}\left(y^*\right) = 0$ and equation (14) would read:

$$\beta_1 = \frac{b_2}{b_1 + \left(\frac{\sigma^2}{\mathrm{var}(y_{t_0}) b_1} \right)} \tag{15}$$

Given inequality (11), $b_1 > b_2$, and the fact that variances are non-negative by definition equation (15) implies a value of β_1 between zero and unity, $0 < \beta_1 < 1$. Consequently, under the assumption that all economies had equal steady state levels of per capita GDP the Solow-Swan model predicts that an OLS-estimation of equation (12) yields $0 < \beta_1 < 1$. A value of $\beta_1 < 0$ or $\beta_1 > 1$ would not support the Solow-Swan model. Hence, although the OLS-estimator of β_1 does no not correspond to $e^{-\beta t_1}$, as in the case where y_{i,t_0} is known, an OLS-estimation of equation (12) yields information, whether the data support the Solow-Swan model or not.

However, while the assumption $\mathrm{var}(y^*) = 0$ may hold for the economies of regions within a country or the economies of countries of similar structure, it is hard to imagine that in holds for a cross section sample of heterogeneous countries. However, given that $\mathrm{var}(y^*) > 0$, equation (14) may yields an OLS-estimator that is upward biased compared to equation (15), because equation (11) implies $a_1 a_2 > a_1^2$, such that the following inequality holds:

$$\frac{a_1 a_2 \frac{\mathrm{var}(y_i^*)}{\mathrm{var}(y_{i,t_0})} + b_1 b_2}{a_1^2 \frac{\mathrm{var}(y_i^*)}{\mathrm{var}(y_{i,t_0})} + b_1^2 + \frac{\sigma^2}{\mathrm{var}(y_{i,t_0})}} > \frac{b_1 b_2}{b_2^2 + \frac{\sigma^2}{\mathrm{var}(y_{t_0})}} \tag{16}$$

Indeed, if $\mathrm{var}(y^*)$ is large compared to $\mathrm{var}(y_0)$ and if σ^2 is not too large, an estimation can even result in a value of $\beta_1 > 1$. Consequently, an OLS-estimation of equation (12) allows no conclusion whether the data support the Solow-Swan model or not, if one can not be sure that $\mathrm{var}(y^*) = 0$ holds. Therefore, Barro/Sala-i-Martin (1992) and Mankiw/Romer/Weil (1993) propose to add proxies for the different steady state levels of y^* into regression equation (12). As the estimation results of β_1 are then „conditioned" on the different steady levels of the economies, they call a value of $\beta_1 < 1$ „conditional beta-convergence". To analyze the effect of this procedure on the OLS-estimator of β_1 consider the following regression equation:

$$y_{i,t_2} = \alpha + \beta_1 \; y_{i,t_1} + \beta_2 \; y_i^* + \mu_i \tag{17}$$

Now the OLS-estimator of β_1 equals:

$$\beta_1 = \frac{\operatorname{cov}(y_{i,t_2}, y_{i,t_1}) \operatorname{var}(y_i^*) - \operatorname{cov}(y_{i,t_2}, y_i^*) \operatorname{cov}(y_{i,t_1}, y_i^*)}{\operatorname{var}(y_{i,t_1}) \operatorname{var}(y_i^*) - \operatorname{cov}(y_{i,t_1}, y_i^*)^2} \tag{18}$$

Inserting equations (9) and (10) and using the relations given in box 1 this can be rewritten to yield again equation (15). Consequently, although equation (9) implies that adding y^* leads to multicollinearity between the explanatory variables y_{i,t_1} and y^*, the Solow-Swan model implies that under this procedure the OLS-estimator of β_1 lies again between zero and unity, $0 < \beta_1 < 1$. Therefore, under this procedure, an OLS-estimation of β_1 yields once more information, whether the data support the Solow-Swan model or not. The econometric explanation for this effect lies in the fact, that the Gauß-Markov theorem implies, that an OLS-estimation yields even in the presence of strong multicollinearity the best, lineary unbiased estimators (BLUE-property of the OLS-estimation) given the standard assumptions of the Gauß-Markov theorem hold. The intuition for this can be grasped by multiplying equation (18) by $1/\operatorname{var}(y^*)$. This shows that the OLS-estimator of β_1 corrects for the multicollinearity by the factor $\operatorname{cov}(y_{i,t_1}, y_i^*)/\operatorname{var}(y^*)$, which equals the coefficient of a single regression between y_{i,t_1} and y^*. However, as is well known, even though an OLS-estimation in the presence of multicollinearity yields a consistent estimation of the regression coefficients, it deflates the standard errors of the regression coefficients. Consequently, an OLS-estimation of equation (16) tends to reduce the level of significance of the estimators. This is however a minor problem, if the estimation yields significant estimation results although the standard errors are deflated.

Equation (16) presents an interesting possibility to empirically test the results of the above analysis. Given a cross section of countries with different steady state variables, equation (16) implies that an OLS-estimation of equation (12) must yield a greater estimation value of β_1 than an OLS-estimation of equation (17). Table 2 presents the results of this test. The data are taken from the Summers/Heston (1991) and the Barro/Lee (1993) data sets. Following the „augmented“ Solow-Swan model of Mankiw/Romer/Weil (1991) the steady state variables of per capita GDP are: the population growth rate, n, the investment quota, s_k, and the secondary school enrollment ratio, s_h. The level of per capita GDP of the base period is approximated by the level of per capita GDP of the year 1960, y_0.

Table 2 - OLS-estimation of equation (12) and (17)[1]

	(Y/L) 1985		(Y/L) 1980		(Y/L) 1975		(Y/L) 1970		(Y/L) 1965	
CONST	0.518	1.076	0.407	0.463	0.422	0.345	0.168	0.061	0.027	-0.059
	(0.16)	(0.007)	(0.188)	(0.000)	(0.06)	(0.157)	(0.259)	(0.000)	(0.791)	(0.648)
y_0	1.001	0.678	1.015	0.787	0.999	0.836	1.018	0.938	1.017	0.979
p-value(β=0)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
p-value(β=1)	-	(0.000)	-	(0.000)	-	(0.000)	-	(0.011)	-	(0.235)
n	-	-0.137	-	-0.089	-	-0.089	-	-0.065	-	-0.03
p-value(β=0)		(0.006)		(0.000)		(0.000)		(0.000)		(0.000)
s_k	-	0.362	-	0.393	-	0.287	-	0.141	-	0.088
p-value(β=0)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
s_h	-	0.023	-	0.012	-	0.01	-	0.005	-	0.001
p-value(β=0)		(0.003)		(0.037)		(0.000)		(0.204)		(0.786)
$\overline{R}^2$	0.799	0.877	0.854	0.906	0.913	0.946	0.962	0.974	0.981	0.984
p-value regr. F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
s.e.e.	0.513	0.39	0.413	0.331	0.31	0.243	0.204	0.169	0.142	0.129
D.W.	2.05	2.195	2.046	2.175	2.158	2.314	2.044	2.114	1.846	1.973
Observations	96	96	97	97	100	100	99	99	99	99

[1] Variables as defined in text. 'p-value(β=0)' = p-value of a test that the of regression coefficient equals zero. 'p-value(β=1)' p-value of a test that the regression coefficient equals one. 'p-value of F' = p-value of an F-test of the null-hypothesis that all regression coefficients are jointly zero. $\overline{R}^2$ = adjusted R-squared value. 'S.e.e.' = Standard error of estimate. 'D.W. ' = Durbin Watson test of serial correlation of the regression errors (The sample is ranked according to the 1985 level of per capita GDP.).

The results of table 2 do not reject the implications of the Solow-Swan model on conditional beta-convergence. First, controlling for the different steady state level of per capita GDP leads indeed to a smaller absolute value of the coefficient than controlling not. This is implied by inequality (16). Second, the coefficient of per capita GDP of the base period is significantly (with exception of 1965) smaller than unity as implied by equality (15), indicating conditional beta-convergence. Third, the coefficient of per capita GDP of the base period increases as the distance between the base period an the period of the level of per capita GDP, which is to be explained, is reduced. This is also implied by equation (15). These results do not significantly change if the data sample is restricted: Table 2 is based on the largest sample of countries for which steady state variables are available. It includes 99 countries. Restricting these countries to the those 93 countries of Barro/Sala-i-Martin (1995), chapter 11 (this sample excludes i.a. all large oil countries), for which the steady state variables are available, yields similar estimates. Hence, the implication of the Solow-Swan model concerning conditional beta-convergence are supported by the data. Next I turn to the implications of the Solow-Swan model concerning sigma-convergence.

4. Conditional and unconditional sigma-convergence: Implications of the Solow-Swan model

In this section I discuss the implications of the Solow-Swan model for an OLS-estimation of sigma-convergence. Following the definition of section 2 sigma-convergence implies that $\mathrm{var}(y_{i,t})$ decreases over time. Following equations (9) and (10) the variance of $y_{i,t}$ at two different points in time, t_1 and t_2 equals:

$$\mathrm{var}(y_{i,t_1}) = a_1^2\ \mathrm{var}(y_i^*) + b_1^2\ \mathrm{var}(y_0) + \sigma^2 \tag{19}$$

$$\mathrm{var}(y_{i,t_2}) = a_2^2\ \mathrm{var}(y_i^*) + b_2^2\ \mathrm{var}(y_0) + \sigma^2 \tag{20}$$

As $a_1 < a_2$ and $b_1 > b_2$, these equations imply sigma-convergence, i.e. $\mathrm{var}(y_{i,t_1}) > \mathrm{var}(y_{i,t_2})$, only if $\mathrm{var}(y_i^*) < \mathrm{var}(y_{i,0})$. Hence, within a cross-section of economies, which have the same steady state parameters, such that $\mathrm{var}(y_i^*) = 0$, the Solow-Swan model predicts sigma-convergence. As the assumption of more or less identical steady state parameters may hold for the economies of the regions within a country, one can draw the conclusion that the Solow-Swan model predicts sigma-convergence for the regions within a country. This implication is supported by the results of (Barro/Sala-i-Martin (1995), chapter 11) for the U.S. States over a long period of time 1890-1990 and by their findings for Japanese prefectures and European regions.

However, given a cross section of countries, where the steady state parameters deviate, such that $\mathrm{var}(y_i^*)>0$, the Solow/Swan model is compatible with sigma-convergence or sigma-divergence. Nevertheless, it is possible to show that the model predicts, what is called „conditional sigma-convergence“ in the following. To see this, consider the following single variable regression:

$$y_{i,t_1} = \alpha + \beta_1\ y_i^* + \mu_{i,t_1} \tag{20}$$

The OLS-estimator of α resp. β_1 equals:

$$\alpha = E(y_{i,t_1}) - \beta_1\ E(y_i^*) \tag{21}$$

$$\beta_1 = \frac{\text{cov}\left(y_{i,t_1}, y_i^*\right)}{\text{var}\left(y_i^*\right)} \tag{22}$$

Consequently, given some of the relations of box 1, based on these formulas the estimated value of y_{i,t_1} based on regression equation (21), $\hat{y}_{i,t_1}$, is given by the following equation:

$$\hat{y}_{i,t_1} = b_1 \, E\left(y_{i,0}\right) + a_1 \, y_i^* \tag{23}$$

Consequently, following equation (9), the residuum of regression equation (20) equals:

$$\mu_i = b_1 \left(y_{i,0} - E\left(y_{i,0}\right)\right) + \varepsilon_{i,t}$$

The variance of this residuum can then be written:

$$\text{var}\left(\mu_{i,t_1}\right) = b_1^2 \; \text{var}\left(y_{i,0}\right) + \sigma^2 \tag{24}$$

Consequently, as $b_1 = e^{-\beta t_1}$, this variance must decrease in the course of time, if the implications of the Solow-Swan model hold. This is called „conditional sigma-convergence".

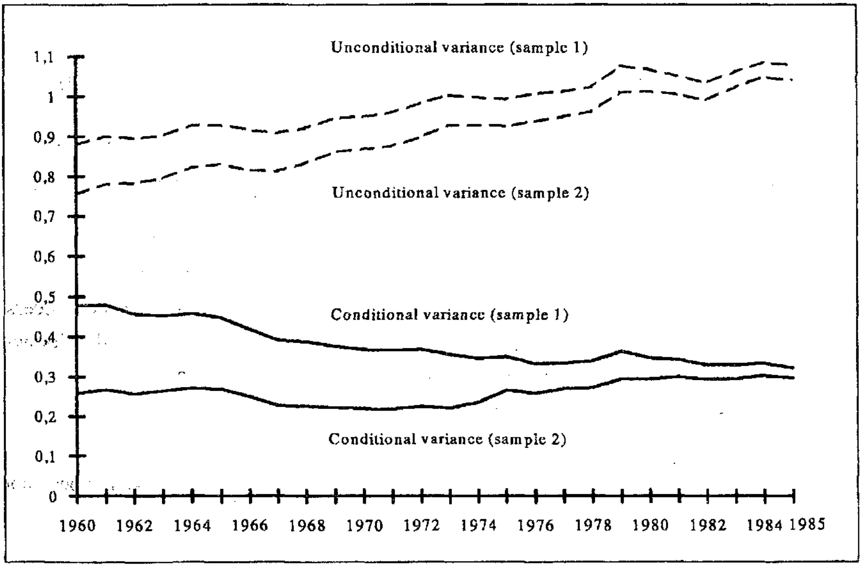
The above derivation of unconditional resp. conditional sigma-convergence depend on the assumption implied by equations (9) and (10) that the stochastic deviations from the level of per capita GDP, which is explained by the Solow-Swan model, are a mean reverting process *and* that a stochastic deviation of per capita GDP from the value explained by the Solow-Swan model is *not* transmitted into the next period. Barro/Sala-i-Martin (1995), pp. 31-32, and pp. 383-386, show that under the assumption that stochastic deviations from the level of per capita GDP are transmitted into future periods (i.e. $y_{i,t} = a - e^{-\beta} y_{i,t-1} + \varepsilon_t$, see Barro/Sala-i-Martin equations (1.20) resp. (11.1)) the Solow-Swan model does not necessarily imply the existence of conditional resp. unconditional sigma-convergence.

Figure 1 displays the results of an estimation of conditional and unconditional sigma-convergence for a cross section of 99 countries of the Summers/Heston (1991) data, sample 1 (this sample includes all the countries for which the steady state variables are available) and a cross section of 93 countries of the same data set, sample 2 (this sample includes all the countries of Barro/Sala-i-Martin (1995), chapter 11, for which the steady state variables are available). The proxies for the steady state variables correspond to those of table 2.[8] As the

[8] As the Barro/Lee (1993) data on human capital refer only to the years 1960,1965,1970,1975,1980,1985, I interpolate between these years to derive human capital proxies for every year.

figure shows, according to the implications of the Solow-Swan model, the conditional variance for this section of countries decreases for sample 1. However, as sample 2 shows, excluding only six countries (Afghanistan, Fiji, Iceland, Kuwait, Mozambique, Myanmar) yields a result that corresponds more to sigma-constancy that sigma-convergence. As sample 2 equals the Barro/Sala-i-Martin (1995) sample this indicates that the finding of conditional-beta convergence for this sample goes probably hand in hand with conditional sigma-constancy. The fact that the unconditional variance increases for both samples indicates than $\mathrm{var}(y_i^*) > \mathrm{var}(y_{i,0})$, i.e. that in steady state this cross section of countries will display a greater variance of the log of per capita income than in the beginning.

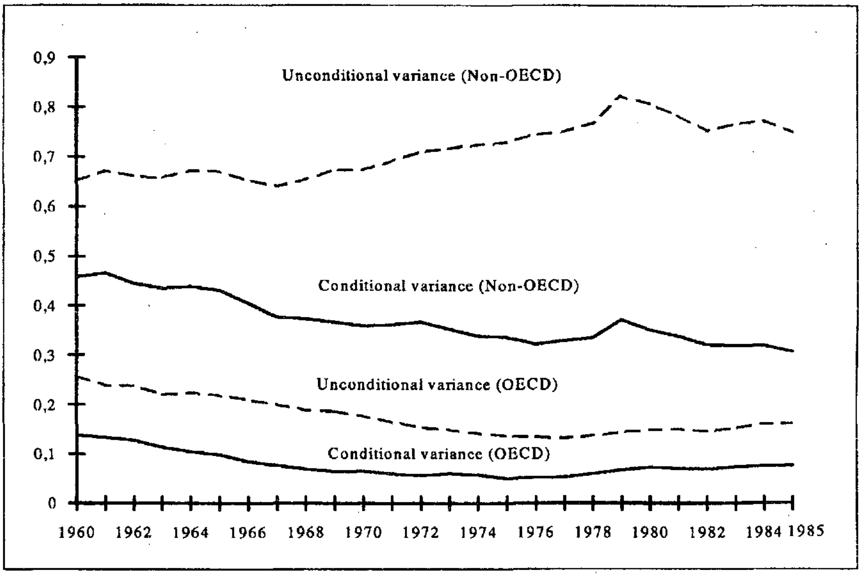
Figure 1 - Unconditional and conditional sigma-convergence across countries (a)



(a) For the definition of sample 1 and 2 see the text above.

Figure 2 displays the results of an estimation of conditional and unconditional sigma-convergence for the sample 1 with the only difference that the OECD-countries were singled out. As this shows, for both subsamples, conditional sigma-convergence holds, as implied by the Solow-Swan model. However, the OECD-countries display also unconditional sigma-convergence. This indicates that the steady state variables of per capita GDP are similar across the OECD-countries, such that they approach to more or less equal levels of per capita GDP and that hence their $\mathrm{var}(y_i^*)$ is small.

Figure 2 - Unconditional and conditional sigma-convergence for OECD countries (a)



(a) For the definition of Non-OECD countries see the text above.

5. Conclusions

This paper has analyzed the implications of the Solow-Swan model for an estimation of conditional and unconditional beta- and sigma-convergence. I argued that given the implications of this model a consistent estimation of conditional and unconditional beta- and sigma-convergence should be possible. The results concerning conditional beta-convergence are not able to reject the implications of the models. However, as implied by lemma 3 and 6 this does not necessarily imply the existence of conditional sigma-convergence: Beta-convergence is compatible with sigma-constancy or even sigma divergence. The results of the tests on conditional sigma-convergence are not that unequivocal. They are sensitive to small alternations of the country sample. However, based on these results of conditional sigma-convergence I would not draw the conclusion that this is serious empirical evidence against the implications of the Solow-Swan model. The existence of *unconditional* sigma-convergence over a long span of time and over such different data sets as the U.S. states, Japanese prefectures or European regions (Barro/Sala-i-Martin (1995), chapter 11) indicates that the non-robust results on conditional sigma-convergence might stem from an insufficient quality of the steady state variables used for estimation.

6. References

BARRO, Robert J., 1991, Economic Growth in a Cross Section of Countries, Quarterly Journal of Economics, Vol. 106, 2, 407-443.

BARRO, Robert J., Xavier, SALA-I-MARTIN, 1991, Convergence across States and Regions, Brookings Papers on Economic Activity, no. 1, 107-182.

BARRO, Robert J., Xavier, SALA-I-MARTIN, 1992, Convergence, Journal of Political Economy, Vol. 100, 2, 223-251.

BARRO, Robert J., Xavier, SALA-I-MARTIN, 1995, Economic Growth, McGraw-Hill Inc. New York.

BARRO, Robert. J., Jong -Wha LEE, 1993, International comparisons of educational attainment, NBER Working Paper No. 4349.

IRTU (International Road Transport Union), 1990, World Transport Data, Centre International, Genf.

MANKIW, N. Gregory, David ROMER, David N.WEIL, 1992, A Contribution to the Empirics of Economic Growth, Quarterly Journal of Economics, Vol. 107, 2, 407-437.

QUAH, Danny, 1993a, Galton's Fallacy and Tests of the Convergence Hypothesis, Scandinavian Journal of Economics, 95, Vol. 4, 427-443.

QUAH, Danny, 1993b, Empirical cross-section dynamics in economic growth, European Economic Review, Vol. 37, pp. 426-434.

SUMMERS, Robert, Alan HESTON,1991, The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988, Quarterly Journal of Economics, pp. 327-368.