University of Victoria
Department of Economics

Econometrics Working Paper EWP0502

ISSN 1485-6441

# Convergence of Income Among Provinces in Canada – An Application of GMM Estimation

**Mukesh Ralhan**
Department of Economics, University of Victoria
*Victoria, B.C., Canada*
**&**

**Ajit Dayanandan**
Department of Economics, University of Northern British Columbia
*Prince George, B.C., Canada*

***March, 2005***

**Abstract**

This paper tests for unconditional and conditional income convergence among provinces in Canada during the period 1981-2001. We apply the first-differenced GMM estimation technique to the dynamic Solow growth model and compare the results with the other panel data approaches such as fixed and random effects. The method used in this paper accounts for not only province-specific initial technology levels but also for the heterogeneity of the technological progress rate between the 'richer' and 'not so richer' provinces of Canada. One of the findings of the paper is that the Canadian provinces do not share a common technology progress rate and a homogeneous production function. The findings of the study suggest a convergence rate of around 6% to 6.5% p.a. whereas the previous studies using OLS and other techniques reported a convergence rate of around 1.05 % for *per capita* GDP and 2.89% p.a. for personal disposable income among Canadian provinces.

**Keywords:** Provincial convergence, Canada, Panel data, GMM

**JEL Classifications:** O40, O41, O53, C14, H4

**Author Contact:**
Mukesh Ralhan, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada, V8W 2Y2; e-mail: mralhan@uvic.ca; FAX: (250) 721-6214

## 1. INTRODUCTION

In recent years, the issue of *per capita* income convergence has been an area of intense research and investigation. The idea of convergence is based on the neoclassical growth model (Solow, 1956, Swan, 1956). Neoclassical theory predicts that if different countries (or provinces) are at different points relative to their balanced growth paths, poorer countries or regions will grow faster than the rich ones, the so-called β (beta) convergence. Evidence of this 'catch-up' effect is interpreted as support for the neo-classical growth model. Various studies have examined the convergence hypothesis in different parts of the world with varying results. The empirical literature distinguishes three distinct, though related, concepts of convergence, see Hossain (2000): sigma (σ) convergence, beta (β) convergence and conditional beta ($\beta^c$) convergence. Another type of convergence, called stochastic convergence, focuses on the time series properties of the distribution of *per capita* income. We explain these terms in the following sections. **Sigma convergence** is concerned with cross-sectional dispersion of *per capita* income or productivity levels; that is, there exists a convergence if the cross-sectional dispersion of *per capita* income or productivity levels decrease over time. The standard deviation of (the log of) *per capita* output is commonly used to test for sigma convergence. Thus the presence of sigma convergence suggests a tendency to equalization of *per capita* income or productivity levels across regions or economies. The presence of sigma convergence, however, does not necessarily imply the presence of beta convergence, which suggests that poorer countries or regions grow at faster rates than richer countries or regions given that they all have the same steady-state growth path for per capita output.

Whether the convergence of *per capita* output levels, measured by sigma convergence, is due to higher growth rates of poorer regions than the richer ones can be examined by testing for the presence of beta (or conditional beta) convergence. **Beta convergence**, as defined in the empirical literature is concerned with cross-section regression of the time averaged income growth rate on the initial *per capita* income level; that is, there exists a beta convergence if the coefficient of the initial *per capita* income level in a cross-section regression for *per capita* output growth bears a negative sign. This suggests that countries or regions with higher initial income levels grow slowly than countries or regions with lower initial income levels. The concept of **conditional beta convergence** concerns with cross-section regression of the time averaged output growth rate on the initial *per capita* output level and a set of additional explanatory variables that define the steady-state growth path for *per capita* output (Barro and Sala-i-Martin, 1991). Within such augmented growth regression, there exists a conditional beta convergence if the coefficient of the initial *per capita* output level bears a negative sign. One of the most generally accepted results is that while there is no evidence of unconditional convergence among a broad sample of countries; the conditional convergence hypothesis holds when examining more homogeneous group of countries (or regions) or when conditioning for additional explanatory variables (Baumol, 1986, Barro and Salai-i-Martin, 1992, and Mankiw *et al.,* 1992 (hereafter MRW)).

For expositional purposes, this paper is divided into five sections: section 2 briefly reviews the methodological framework; section 3 touches upon the data issue; section 4 is devoted to empirical investigation and section 5 summarizes the conclusions emanating from the study.

## 2. METHODOLOGY

The issue of the presence and persistence of regional disparities in income and output *per capita* in Canada has been a thorny issue and has concerned both economists and politicians. Estimates of convergence (β convergence) over a wide range of data suggest a convergence rate that ranges from 1.05% p.a. for gross provincial product *per capita* to 2.89 % p.a. for personal disposable income *per capita* (see, for example, Coulombe and Lee (1995) and Lee and Coulombe (1995)). Lee and Coulombe (1995) have applied the OLS method in a pooled regression for their estimates. Kaufman *et al.* (2003) recently studied the differential impact of federal transfer programs on output convergence in provinces in Canada using three stages least squares method. They found that while the employment insurance (EI) system seems to have had a significant negative effect on output convergence – by discouraging migration within Canada, the equalization transfers might have actually helped spur convergence. In yet another study Wakerly (2002), using data on provinces and industries found that the process of economic growth in Canada has resulted in poor provinces staying poor and the rich provinces staying rich. These contrasting conclusions raise questions about the divergent methodologies used by researchers in analyzing this issue in the Canadian context. One of the major limitations of studies on convergence of income in Canadian provinces is the absence of any studies that addresses this issue using panel data methods or techniques. The main usefulness of the panel approach lies in its ability to allow for differences in the aggregate production function across economies. This leads to results that are significantly different from those obtained from single cross-country regressions (Islam, 1995).

Thus, the main aim of this paper is first, to examine the issue of income convergence in Canada using sound methodologies such as panel data methods and/or Generalized Method of Moments (GMM) estimation techniques, and to see how the results thus obtained differ from other estimations. If the findings do suggest widening gap or slow regional income convergence, another objective is to stimulate the debate on policy formulations in regard to balanced regional development in Canada. We apply the Fixed Effect (FE) and Random Effect (RE) techniques and compare our results with first-differenced GMM as the Solow growth model that we estimate in our study is a dynamic model and hence the endogeneity issue has to be addressed. Another method that could be applied is the system GMM estimator developed by Blundell and Bond (1998) that presents a significant improvement over the other panel estimation methods such as the first-difference GMM estimator (see Arellano and Bond (1991)). Using Monte Carlo simulations, they demonstrate that the weak instruments problem in first-difference GMM could result in large finite-sample biases. Also the biases can be dramatically reduced by incorporating more informative moment conditions that are valid under quite reasonable stationarity restrictions on the initial condition process. Essentially the system GMM estimator represents the use of lagged first-differences as instruments for equations in levels, in addition to the usual lagged levels as instruments for equations in first-differences. In addition, the finite-sample performance of the system GMM can be tested by the identification of an estimation range for the convergence speed provided by the OLS and other estimators like within-group estimator. In this paper, however, we confine our study to the application of the Fixed and Random Effects and the first-differenced GMM estimators. The application of system GMM could be considered for future research work.

With a view to conserving space, our intention is not to review the basic Solow growth model in detail here. See studies by MRW (1992) and Islam (1995) that give a detailed description of the basic Solow growth model and augmented Solow growth model. In the following sections, we succinctly explain the usefulness of a panel data approach on the basis of the work by Islam (1995). The basic Solow model is discussed briefly below (MRW, 1992):

The rates of saving, population growth and technological progress are treated as exogenous. There are two inputs, capital and labor, which are paid their marginal products. Solow assumes a Cobb-Douglas production function, so production at time t is given by:

$$Y(t) = K(t)^{\alpha} (A(t)L(t))^{1-\alpha} \quad ; \quad 0 < \alpha < 1 ,$$

where $Y$ is output, $K$ is capital, $L$ is labor, and $A$ is the level of technology. $L$ and $A$ are assumed to grow exogenously at the rates $n$ and $g$:

$$L(t) = L(0)\, e^{nt}$$

$$A(t) = A(0)\, e^{gt} \quad .$$

Assuming that $s$ is the constant fraction of output that is saved and invested, and defining output and stock of capital per unit of effective labor as $\hat{y} = Y / AL$ and $\hat{k} = K / AL$, respectively, the dynamic equation for $\hat{k}$ is given by:

$$\dot{\hat{k}}(t) = s\,\hat{y}(t) - (n + g + \delta)\,\hat{k}(t)$$

$$= s\,\hat{k}(t)^{\alpha} - (n + g + \delta)\,\hat{k}(t)$$

where $\delta$ is the constant rate of depreciation. It is evident that $\hat{k}$ converges to its steady state value:

$$\hat{k}^{*} = \left(\frac{s}{n + g + \delta}\right)^{1/(1-\alpha)} .$$

Upon substitution this gives the following expression for steady state *per capita* income:

$$\ln\left[\frac{Y(t)}{L(t)}\right] = \ln A(0) + gt + \frac{\alpha}{1-\alpha}\ln(s) - \frac{\alpha}{1-\alpha}\ln(n + g + \delta) \qquad \textbf{(1)}$$

Assuming that the countries are currently in their steady state, MRW (1992) used equation (1) to see how differing saving and labor force growth rates can explain the differences in the current *per capita* incomes across countries. They found the model to be quite successful except that the estimates of the elasticity of output with respect to capital, $\alpha$, were found to be unusually high. In order to overcome this problem, they suggested augmented Solow growth model with human capital as another input of the production and hence as a variable in the regression equation. MRW assumed that $g$ and $\delta$ are constant across countries. But the $A(0)$ term in equation (1) reflects not just technology but resource endowments, climate, institutions, *etc*.. It may therefore differ across countries. So they assumed that $Ln\ A(0) = a + \varepsilon$, where $a$ is a constant and $\varepsilon$ is a country-specific shock. Thus log income *per capita* at a given time – time 0 for simplicity is:

$$\ln\left[\frac{Y}{L}\right] = a + \frac{\alpha}{1-\alpha}\ln(s) - \frac{\alpha}{1-\alpha}\ln(n+g+\delta) + \varepsilon \tag{2}$$

At this stage, however, MRW made the assumption that $\varepsilon$ is independent of the explanatory variables, $s$ and $n$. This was their identifying assumption, and it allowed them to proceed with the Ordinary Least Squares (OLS) estimation of the equation.

Another specification used by MRW is to include human capital in the production function in view of the importance of human capital to the process of growth. Therefore, the production function now becomes:

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

where $H$ is 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 of income invested in human capital. The evolution of the economy is now determined by:

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

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

where $y = (Y/AL)$, $k = (K/AL)$, and $h = (H/AL)$ are quantities per effective unit of labor, and $(\alpha + \beta) < 1$, which implies that there are decreasing returns to all capital. The economy converges to a steady state defined by:

$$k^* = \left(\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta}\right)^{1/(1-\alpha-\beta)} \quad ; \quad h^* = \left(\frac{s_k^{\alpha} s_h^{1-\alpha}}{n+g+\delta}\right)^{1/(1-\alpha-\beta)}$$

Substituting $k^*$ and $h^*$ into the production function and taking logarithms gives an equation for income *per capita* similar to equation (1):

$$\ln\left[\frac{Y(t)}{L(t)}\right] = \ln A(0) + gt - \frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \frac{\beta}{1-\alpha-\beta}\ln(s_h) \qquad \textbf{(3)}$$

MRW estimated equation (3) using the OLS method. While estimating equation (2), MRW assumed that $\varepsilon$ is independent of the explanatory variables, $s$ and $n$. However, Islam (1995) argues that in general, the country-specific technology shift term $\varepsilon$ is likely to be correlated with the saving and population growth rates experienced by that country. At a heuristic level, as $A(0)$ is defined not only in the narrow sense of production technology, but also to include resource endowments, institutions, *etc*., it is not entirely convincing to argue that saving and fertility behavior will not be affected by all that is included in $A(0)$. What is important to note here is that in the framework of a single cross-section regression, this assumption of independence becomes an econometric necessity. OLS estimation is valid only under this assumption.

For the above reasons, Islam (1995) proposes that a panel data framework provides a better and more natural setting to control for this technology shift term $\varepsilon$. Islam (1995) has derived the steady state behavior in the following manner:

Let $\hat{y}^*$ be the steady state level of income per effective worker, and let $\hat{y}(t)$ be its actual value at any time t. Approximating around the steady state, the rate of convergence is given by:

$$\frac{d\ln\hat{y}(t)}{dt} = \lambda[\ln\hat{y}^*) - \ln\hat{y}(t)],$$

where $\lambda = (n+g+\delta)(1-\alpha)$.

This equation implies that:

$$\ln\hat{y}(t_2) = (1-e^{-\lambda\tau})\ln\hat{y}^* + e^{-\lambda\tau}\ln\hat{y}(t_1) \ ,$$

where $\hat{y}(t_1)$ is income per effective worker at some initial point of time and $\tau = (t_2 - t_1)$. Subtracting $\hat{y}(t_1)$ from both sides yields

$$\ln\hat{y}(t_2) - \ln\hat{y}(t_1) = (1-e^{-\lambda\tau})\ln\hat{y}^* - (1-e^{-\lambda\tau})\ln\hat{y}(t_1) \, .$$

This equation represents a partial adjustment process that becomes more apparent from the following rearrangement.

$$\ln \hat{y}(t_2) - \ln \hat{y}(t_1) = (1 - e^{-\lambda\tau})(\ln \hat{y}^* - \ln \hat{y}(t_1)\,.$$

In the standard partial adjustment model, the "optimal" or "target" value of the dependent variable is determined by the explanatory variables of the current period. In the present case, $\hat{y}^*$ is determined by *s* and *n*, which are assumed to be constant for the entire intervening time period between $t_1$ and $t_2$ and hence represent the values for the current year as well. Substituting for $\hat{y}^*$ gives:

$$\begin{aligned} &\ln \hat{y}(t_2) - \ln \hat{y}(t_1) = (1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s) \\ &- (1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n + g + \delta) - (1 - e^{-\lambda\tau})\ln \hat{y}(t_1) \end{aligned} \quad . \qquad \textbf{(4)}$$

MRW used this equation to study the process of convergence across different samples of countries. Islam (1995) reformulated the equation in terms of income per capita. Note that income per effective labor is:

$$\hat{y}(t) = \frac{Y(t)}{A(t)L(t)} = \frac{Y(t)}{L(t)A(t)e^{gt}}\,,$$

so that

$$\begin{aligned} \hat{y}(t) &= \ln\left(\frac{Y(t)}{L(t)}\right) - \ln A(0) - gt \\ &= \ln y(t) - \ln A(0) - gt \end{aligned}$$

where *y*(*t*) is the *per capita* income, [*Y*(*t*)/*L*(*t*)]. Substituting for $\hat{y}$ into the above equation, we get the usual "growth-initial level" equation:

$$\ln y(t_2) - \ln y(t_1) = (1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s)$$
$$-(1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n + g + \delta) - (1 - e^{-\lambda\tau})\ln y(t_1) + (1 - e^{-\lambda\tau})\ln A(0) + g(t_2 - e^{-\lambda\tau}t_2)$$

However, if we collect terms with *ln* $y(t_1)$ on the right-hand side, we can write the equation in the following alternative form:

$$\ln y(t_2) - \ln y(t_1) = (1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s)$$
$$-(1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n + g + \delta) + e^{-\lambda\tau}\ln y(t_1) + (1 - e^{-\lambda\tau})\ln A(0) + g(t_2 - e^{-\lambda\tau}t_2) \quad \textbf{(5)}$$

It can now be seen that the above represents a dynamic panel data model with $(1\text{-}e^{\lambda\tau})ln\ A(0)$ as the time-invariant individual country-effect term. We may use the following conventional notation of the panel data literature:

$$y_{it} = \gamma y_{i,t-1} + \sum_{j=1}^{2}\beta_j x_{it}^j + \eta_t + u_i + v_{it}, \quad \textbf{(6)}$$

where

$$y_{it} = \ln y(t_2)$$
$$y_{i,t-1} = \ln y(t_1)$$
$$\gamma = e^{-\lambda\tau},\ \ \beta_1 = (1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha},\ \ \beta_2 = -(1 - e^{-\lambda\tau})\frac{\alpha}{1-\alpha}$$
$$x_{it}^1 = \ln(s)$$
$$x_{it}^2 = \ln(n + g + \delta)$$
$$u_i = (1 - e^{-\lambda\tau})\ln A(0)$$
$$\eta_i = g(t_2 - e^{-\lambda\tau})t_1 .$$

Panel data estimation of this equation now allows us to control for the individual country effects. Islam (1995) applied Least Squares with Dummy Variable (LSDV) and Minimum Distance (MD) estimation techniques for the estimation of equation (6).

We start our empirical investigations in Section 3, using Fixed and Random Effects techniques and proceed to the first-difference GMM estimator (see Arellano and Bond (1991). Some of the advantages of using the panel data approach are: availability of large number of data points help in reducing problem of limited degrees of freedom and enhance the efficiency of the estimators; problem of multicollinearity is less likely since explanatory variables vary in two dimensions; enables researches to address questions by studying intertemporal behavior of cross-sections, which is otherwise not possible to answer; problem of omitted variables can be reduced by explicitly modeling unobserved variables as a unit-specific effect or time-specific effect or both.

In addition to the fixed and random effects and first-differenced GMM estimators that we discuss in the next section, another approach for dynamic panel data models is the system GMM as suggested by Blundell and Bond (1998). Weeks and Yao (2002) have recently applied the system GMM estimation technique for testing conditional income convergence in provinces of China. They have examined the issue of income convergence in the basic Solow Growth framework. Using a different notation, Weeks and Yao (2002) estimate the equation (6) in the following form:

$$y_{it} = by_{i,t-1} + \theta' x_{i,t}^{j} + T_t + \eta_i + v_{it} \qquad \textbf{(7)}$$

where

$$x_{it} = (\ln(s_{it}), \ln(\eta_{it} + g + \delta)', \theta = \left((1-\varsigma)\frac{\alpha}{1-\alpha}, -(1-\varsigma)\frac{\alpha}{1-\alpha}\right)',$$

$$\eta_i = (1-\varsigma)\ln A(0), \; T_t = g(t_2 - \zeta t_1),$$

and

$$b = 1+\beta = \varsigma..$$

They interpret the effects $\eta_i$ as a composite of unobservable province-specific factors, which includes initial technology differences. Similarly, $T_t$ captures the time-specific effects, which includes the rate of technological change. In order to test that the technology progress rate of coastal provinces is different from that of the interior provinces in China, Weeks and Yao (2002) include a composite dummy constructed by taking the product of a time and coastal dummy ($D_i$). So, they re-write equation (7) as:

$$y_{i,t} = by_{i,t-1} + \theta x_{i,t} + D_i T_t + T_t + \eta_t + v_{i,t} \qquad \textbf{(8)}$$

Following reforms in China, the coastal regions attained a faster rate of growth compared to the interior regions. In order to differentiate between the two regions, Weeks and Yao (2002) have used the coast dummy variable. If the provinces belong to the coastal regions, the dummy variable is equal to 1, and 0 otherwise. In this paper, however, we report results from the fixed effects, random effects and first-differenced GMM estimation techniques and application of System GMM approach will be undertaken for future research work.

## 3. DATA ISSUES

We have used the annual per capita real net provincial domestic product (NPDP) for the period 1981-2001 for 10 provinces *viz*., Newfoundland (NF), Prince Edward Island (PEI), Nova Scotia (NS), New Brunswick (NB), Quebec (QB), Ontario (ON), Manitoba (MB), Saskatoon (SK), Alberta (AB) and British Columbia (BC). Our main source of data is the online database of Statistics Canada: CANSIM II. The variables for which we collected data include: NPDP, labor force growth rate for working age population in the age group 15-64 years and Real Investment. Our empirical study is confined to estimation of the basic Solow growth model and endogenous growth model at this stage. Data sources and exact definitions of variables are available from the authors on request. Following Islam (1995), we use five-year time intervals for averaging the data. By adopting this approach our results are less likely to be influenced by business cycle fluctuations.

In this empirical investigation, we have classified the provinces into (a) Below Average Provinces (Nova Scotia, New Brunswick, Prince Edward Island and Newfoundland, Quebec, Saskatchewan and Manitoba) and (c) Above Average Provinces (Alberta, British Columbia and Ontario) based on their relative performance. By such classification, we would like to test that convergence hypothesis holds more strongly for a homogenous group of countries (or regions) (Barro and Salai-i-Martin, 1992, and Mankiw *et al*., 1992).

Figure 1:

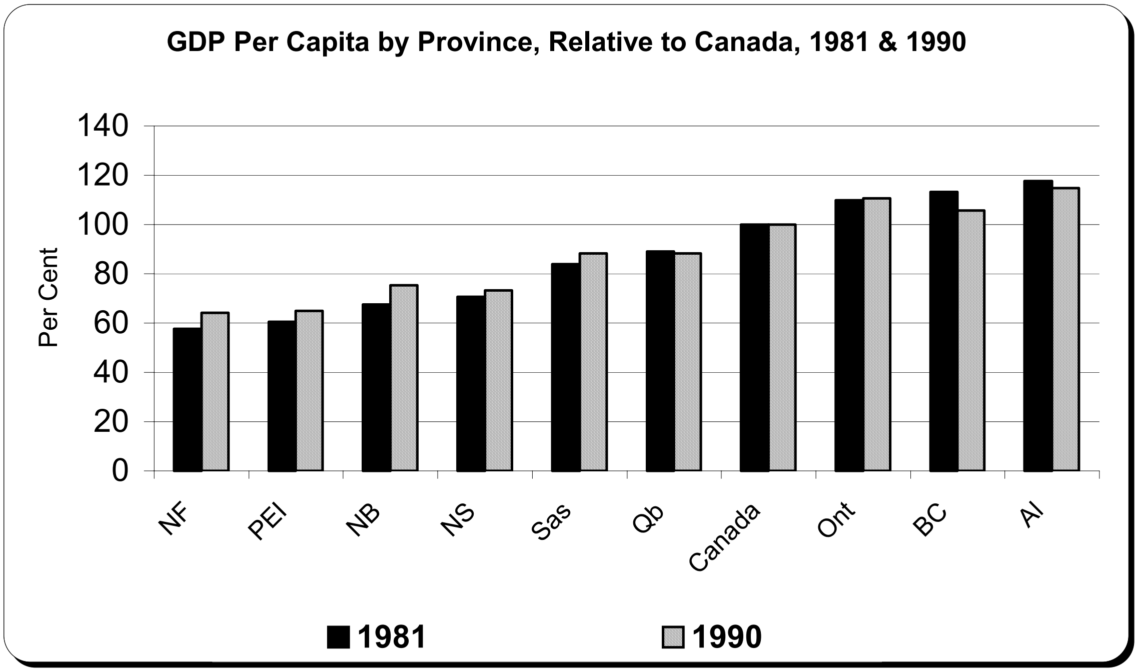



Figure 2:

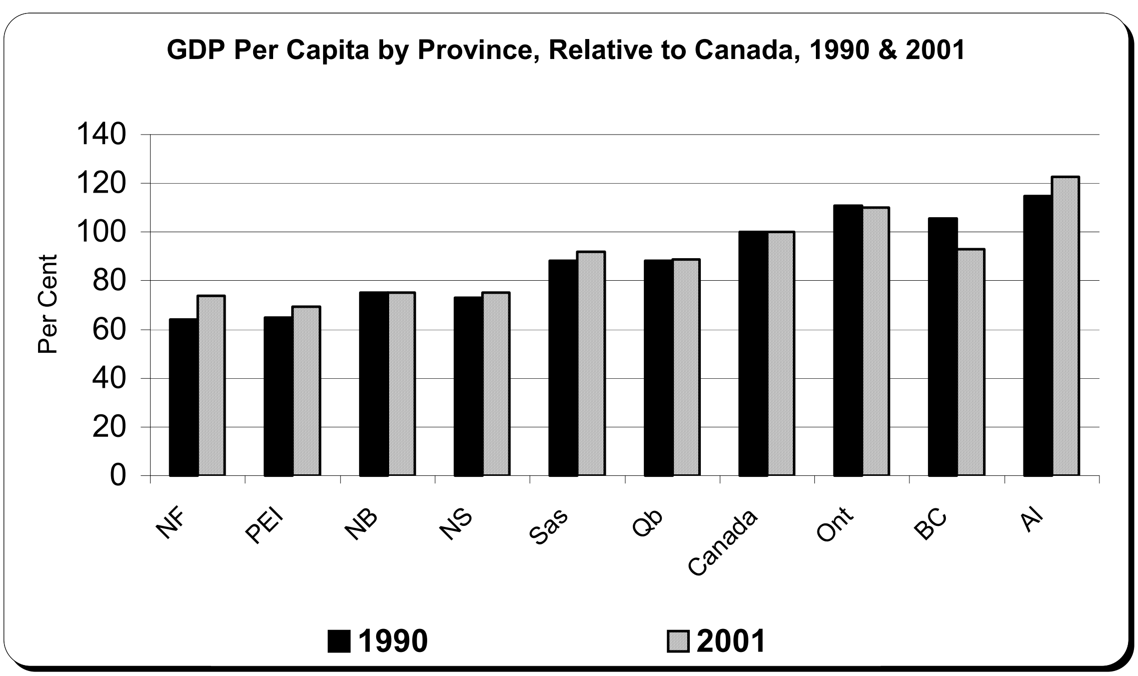


Figure 1 shows the GDP *per capita* by province relative to the Canadian average in 1981 and 1990. There are two striking conclusions: first, there is considerable difference between the average income of the richest province and that of the poorest. Per capita real provincial gross domestic product in Newfoundland in 1990 is only 62 *per cent* of the national average, while in Ontario it was 116 per cent, more than twice as great. By 2001, this picture has marginally changed with the position of provinces like BC showing a marginal deterioration and that of Alberta showing substantial improvement over the national average. Second, these disparities have not changed much in the 1990's as is evident from Figure 2.

## 3. EMPIRICAL INVESTIGATION

In this section, we start our empirical investigation with the endogenous growth model (Barro, 1991 and Barro and Sala-i-Martin, 1991, 1992) used by Coloumbe and Lee (1995). They examined six different concepts of *per capita* income and output convergence using OLS in a pooled regression. Coulombe and Lee (1995) have used the following model in their estimation approach:

$$\frac{1}{10}.\ln\left(\frac{Y_{i,t+10}/\bar{Y}_{t+10}}{Y_{it}/\bar{Y}_{t}}\right) = B - \left(\frac{1-e^{-\beta 10}}{10}\right).\ln\left(\frac{Y_{it}}{\bar{Y}_{t}}\right) + u_{it} \quad , \qquad \textbf{(9)}$$

where $i$ = 1,….,10 (regional units for ten provinces); $t$ = 1961, 1971, 1981, and where $\bar{Y}_t$ refers to the Canadian average (weighted by population) income, $Y$ is output (or income) *per capita*, $B$ is a constant term, u is an error term, and $t$ and T are the initial and the final year of comparison. So, $T$-$t$ is the observation period. Thus Coulombe and Lee (1995) divided the 1961 to 1991 observation period into three sub periods: 1961-71, 1971-81 and 1981-91. Our endeavor would be to apply panel estimation approach to estimating equation (9), which is a test for unconditional convergence hypothesis. We also test for conditional beta convergence as in equation (6) above:

$$y_{it} = \gamma y_{i.t-1} + \sum_{j=1}^{2}\beta_j x_{it}^{j} + \eta_t + u_i + v_{it}\,, \qquad \textbf{(10)}$$

These estimators are called by various names like Least Squares Dummy Variable (LSDV) or fixed effects estimators. $\hat{\beta}$ is unbiased. It is also consistent when either $N$ or $T$ or both tend to infinity. In our case, the number of units, *i.e.* provinces. is equal to ten. However, when the number of units is large, the fixed effects model has too many parameters. The loss of degrees of freedom could be avoided by assuming $u_i$ to be random. Equation (10) could therefore, be rewritten in the following form:

$$y_{it} = \mu + \gamma y_{i,t-1} + \sum_{j=1}^{2} \beta_j x_{it}^j + \eta_t + v_{it}, \qquad \textbf{(11)}$$

where $v_{it} = \alpha_i + u_{it}$. This model is called the random effects (RE) model or error-components model. When the sample size is large, RE (or FGLS) will have the same asymptotic efficiency as GLS. Even for moderate sample size (*e.g.*, $T \geq 3$, N- (K+1) ≥ 9; for $T = 2$, N- (K+1) ≥ 10), the FGLS or RE is more efficient than FE estimator. However, if we are interested in province-specific effect, FE is appropriate. We estimate our model using both the estimators and then apply the Hausman test to determine as to which estimator we should prefer. Another related but important issue is that the model we estimate is a dynamic model and standard estimators like OLS, FE and RE are biased or inconsistent because regressors are correlated with the error term. The consistency of the FE estimator depends on $T$ being large. For RE, the transformed regressors will be correlated with the transformed errors. In order to overcome this problem, Arellano and Bond (1991) suggest that additional instruments can be obtained. They derive a consistent estimator when $N$ goes to infinity with $T$ fixed. Arellano and Bond's preliminary one-step consistent estimator is given by the GLS estimator and then they obtain the optimal GMM estimator. The ultimate resulting estimator is the two-step estimator, which is consistent.

This paper contributes to the existing studies on convergence in Canadian provinces in many ways. First of all, we use more recent time horizon (1981-2001) compared to 1961 to 1991 used by Coulombe and Lee (1995). Second, Coulombe and Lee (1995) have used the OLS method that does not take into account variables that are unobservable and specific to the unit but are time-invariant. Third, the assumption in previous studies is that all the provinces in Canada share a homogenous production function. This is because OLS cannot be applied if the production function is heterogeneous. This paper also contributes in terms of better methodological framework like fixed or random effects and first-differenced GMM.

One of the common methods of measuring persistence is to calculate the half-life[1] of income deviations; *i.e.*, the amount of time it takes a shock to a series to revert half-way back to its mean value. The approximate half-life of a shock to $Y_{it}$ is computed as $-\ln(2)/\ln(\rho_i)$, where $\beta_i \equiv \rho_i - 1$. The persistence parameters $\rho_i$(s) capture the speed of relative income convergence across provinces. Our primary focus is on the $\beta_{is}$, the coefficients on the lagged log of the gross domestic product ($Y_{it}$); the nearer $\beta_i$ is to zero, the longer is the estimated half-life of a shock.

The OLS estimates of $\hat{\rho}$ are downward biased in small samples (Kendall, 1981). In order to correct for the small sample bias, we follow the popular method of adjustment recommended by Nickel (1981) for adjustment of the $\hat{\rho}$ values. The estimated bias-adjusted $\rho$, along with the approximate half-life calculations[2] are reported in Tables (2) and (3). The estimated half-life values give us an idea about the time required (in years) for a province to reach the steady state income level based on the estimated speed of convergence. We compare our results with the speed of convergence estimates obtained by Coulombe and Lee (1995). Table 1 shows the speed of convergence for OECD countries including Canada. The convergence rate for Canada (based on Coulombe and Lee estimation) varies from 1.05% p.a. to 2.89% *p.a.* with estimated half-life from 29 years to 66 years. For other developed countries, it ranges from a low of 0.2% for Denmark to a high of 4.96% *p.a*. for the Netherlands. The lower the speed of convergence, the more time it takes for an economy or province to converge to its steady state equilibrium (measured in half-life). Our objective in the paper is to ascertain whether speed of convergence has undergone a significant change among Canadian provinces ever since Coulombe and Lee (1995) derived their estimates.

Our estimation work is based on two samples. We used the original panel data for the period 1981 to 2001 and the 5-year averaged data for the same period and that resulted in $T = 4$ (*viz*. 1985, 1990, 1995 and 2000) for the ten provinces included in our analysis. We then applied least squares, fixed and random effects and first-differenced GMM estimators to both the samples. Our estimation results are tabulated below in Table 2 and Table 3. We used a number of specifications to test for the convergence hypothesis. However, our basic model continued to be Solow Growth model. In order to compare our results with the Coulombe and Lee (1995) specification, we also tested for unconditional convergence (Table 2). We applied both least squares and panel estimation techniques to see how our results differ from Coulombe and Lee (1995). As mentioned earlier, we classified the provinces into two categories: Above average and below average. For this purpose, we included province-specific dummies to capture the speed of convergence for the respective categories. If the provinces belong to above average category like Alberta, British Columbia and Ontario, the dummy is equal to 1 and 0 otherwise. Table 3 presents results from this specification. Also included are estimation results from the specification which included both time dummies and province-specific dummies.

| **Table 1 – Convergence in OECD countries** | | | | | |
|---|---|---|---|---|---|
| **Unconditional convergence in Canada, 1961-91 (Coulombe and Lee, 1995)** | | | | **Unconditional convergence in other countries (Barro et al, 1991, 1992)** | |
| **Variables** | **β (Rate of Convergence)** | **$R^2$** | **Half-life (years)** | **Country** | **Rate of Convergence** |
| GPP (PRO) | -0.0105 (1.05%) | 0.11 | 66 | U.S. (Unconditional)<br>U.S. (Conditional) | 1.8%<br>2.22% |
| GPP (NAT) | -0.0184 (1.84%) | 0.20 | 38 | Netherlands | 4.96% |
| EI | -0.0162 (1.62%) | 0.21 | 43 | U.K. | 3.37% |
| PIT<br>PI | -0.0163 (1.63%)<br>-0.0241 (2.41%) | 0.18<br>0.29 | 43<br>29 | Belgium | 2.37% |
| PDI | -0.0289 (2.89%) | 0.32 | 24 | Germany<br>Italy<br>France<br>Denmark | 2.30%<br>1.18%<br>1.0%<br>0.2% |

**Source:** Coulombe and Lee (1995)

**Table 2 – Estimates of convergence among provinces in Canada (2004)**

| | **1981-2001** | | | | | **5-year average (1985, 1990, 1995,2000)** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **β** | | **$R^2$** | **Half-life (years)** | | **$\beta^2$** | | **$R^2$** | **Half-life (years)** |
| **Unconditional convergence** | | | | | | | | | |
| Least Squares | -0.0645*** (5.056) | | 0.11 | **10** | Least Squares | -0.0609 * (2.32) | | 0.12 | **11** |
| FE | -0.223[1] *** (10.286) | | 0.35 | **3** | FE | -0.1966[6] *** (3.929) | | 0.48 | **3** |
| RE | -0.0645 ** (3.029) | | 0.11 | **10** | RE | -0.0609 * (2.46) | | 0.12 | **11** |
| GMM (FD) | -0.317[6] *** (16.254) | | 0.29 | **2** | GMM (FD) | -0.387[6] ** (2.684) | | 0.47 | **1.5** |
| **Conditional convergence (Solow Growth Model) (With *n*+*g*+*δ* and Real Savings as additional independent variables)** | | | | | | | | | |
| Least Squares | -0.0630*** (4.70) | | 0.11 | **11** | Least Squares | -0.0617* (2.23) | | 0.15 | **11** |
| FE | -0.223[6] *** (9.818) | | 0.35 | **3** | FE | -0.212[6] *** (4.767 | | 0.47 | **3** |
| RE | -0.0630 *** (4.541) | | 0.11 | **11** | RE | -0.0617 * (2.409) | | 0.15 | **11** |
| GMM (FD) | -0.314[6] *** (10.5) | | 0.2 | **2** | GMM (FD) | -0.471[6] (0.119) | | - | **1.1** |
| Conditional convergence (Solow Growth Model) (With n+g+δ and Real Savings and province dummy as additional independent variables) | | | | | | | | | |
| | β | Dummy | $R^2$ | Half-life (years) | | $\beta^2$ | Dummy | $R^2$ | Half-life (years) |
| Least Squares | -0.137*** (7.324) | 0.005*** (5.335) | 0.22 | **5** | Least Squares | -0.135** (3.51) | 0.005* (2.57) | 0.28 | **5** |
| FE | -0.23*** (9.82) | Dropped | 0.35 | **3** | FE | -0.250 (4.37) | Dropped | 0.44 | **2.5** |
| RE | - 0.137*** (8.670) | 0.005*** (7.074) | 0.22 | **5** | RE | - 0.135** (3.53) | 0.005* (2.582) | 0.28 | **5** |
| GMM (FD) | Near singular matrix | | | | GMM (FD) | Near singular matrix | | | - |

**Notes:** Figures in the parentheses are t-ratios)
*** Significant at 1% level of significance
** Significant at 5% level of significance
*Significant at 10% level of significance

| **Table 3: 5-year average - Conditional convergence (Solow Growth Model) (With $n$+$g$+$\delta$ and real ravings, time dummies and province-specific dummy as additional independent variables)** | | | | |
|---|---|---|---|---|
| **Variables** | **Least Squares** | **Fixed Effect** | **Random Effect** | **First-differenced GMM** |
| Log(GDP)(-1) | -0.114*<br>(2.414) | -0.877***<br>(8.04) | -0.114**<br>(3.292) | -3.104<br>(0.279) |
| Dummy (DT) | .004*<br>(1.976) | - | 0.004***<br>(5.123) | - |
| D90 | -0.0005<br>(0.341) | 0.010***<br>(6.378) | 0.0005<br>(0.667) | 0.0269<br>(0.311) |
| D95 | -0.003<br>(1.841) | -0.010***<br>(4.425) | -0.003<br>(1.688) | 0.548<br>(0.241) |
| D2000 | -0.001<br>(0.512) | 0.020***<br>(6.124) | -0.001<br>(0.352) | 0.092<br>(0.251) |
| Log(n+g+d) | -0.002*<br>(1.960) | -0.0002<br>(-0.913) | -0.002***<br>(4.381) | -0.003<br>(0.121) |
| Log(real saving) | -0.003<br>(0.787) | -0.004<br>(1.051) | -0.003<br>(1.294) | 0.163<br>(0.175) |
| Constant | -.008<br>(1.451) | -0.029***<br>(20.13) | -0.008<br>(1.72) | - |
| $R^2$ | 0.40 | 0.94 | 0.40 | - |
| Half-life (years) | **6** | **0.33** | **6** | **-** |

**Note:** Figures in the parenthesis are t-ratios)
*** Significant at 1% level of significance
** Significant at 5% level of significance
*Significant at 10% level of significance

The null hypothesis of all province-specific effect being the same is strongly rejected in our fixed effects model for the period 1981 to 2001. However, the same is not true for the 5-year average sample. So we don't get support for the use of fixed effects model for the latter. For the random effects model, our null hypothesis is that the regressors and the error term are not correlated. Again, we fail to strongly reject the null hypothesis for the 5-year average sample. In other words, it justifies the use of random effects method for this sample.

We also conducted the Hausman specification test to see which model we should prefer. The test is based on the idea that under the hypothesis of no correlation, both OLS in the LSDV model and GLS are consistent, but OLS is inefficient, whereas under the alternative, OLS is consistent, but GLS is not. Therefore, under the null hypothesis, the two estimates should not differ systematically, and a test can be based on the difference. For the first sample size i.e. 1981 to 2001, the test statistic is greater than the critical value at any level of significance. The hypothesis that the individual effects are uncorrelated with the other regressors in the model, is therefore, strongly rejected. So the use of random effects for this sample is not appropriate. Interestingly, the results from Hausman test for the second sample *i.e*. 5-year average, also provides support for the fixed effects model. However, the problem with the fixed effects model is that it "sweeps out" the effect of province-specific dummy variable. In order to determine these effects, we will have to use the random effects estimator.

Some of the important highlights of our results are as follow:

- Rate of convergence turns out to be much higher than those reported by previous studies. On an average it ranges from 6%% to 6.5% with half-life estimates of around 1 year to 31 years. Our results are consistent with those obtained by Islam (1995). Their findings suggest that panel data estimation techniques give us a higher rate of convergence compared to Least Squares.
- The empirical results from the least squares and random effects estimators are quite close to each other.
- Fixed Effects and First-differenced GMM estimators suggest a much higher rate of convergence.
- Province-specific dummy turns out to be significant in both the specifications. In other words, it suggests that if the province belongs to the above average category, it attains a faster rate of convergence.

- Time-dummies turn out to be insignificant. So the role of time variable towards convergence can be ruled out.
- We also estimated equation (8) above by including both $T_t$ and $DiT_t$ (Table 3). To recapitulate, $T_t$ captures the time-specific effects, which includes the rate of technological change. $DiT_t$ is a composite dummy constructed by taking the product of a time and province dummy. In particular, if we are interested in knowing if the technology progress rate of above average provinces is different from that of the below average provinces, we can test this hypothesis by including composite dummy. This equation gives us a rate of 9.85% (with estimated half-life of around 7 years).

## 4. CONCLUSIONS

Under the Constitution Act of 1982, both federal government and provincial governments are committed to balanced regional development in Canada. The faster rate of convergence of around 6.5% is probably an indication that the programs like federal equalization transfers to provinces and other measures are paying off. If the estimation of convergence rate of around 6.5% among Canadian provinces turns out to be true, then it surpasses the rate reported for U.S. states by a wide margin (which is between 1.8% to 2.22% *p.a.*). The estimated rate of convergence by Coulombe and Lee (1995) related to the time period from 1961 to 1991 whereas the sample period used in our study is more recent i.e. 1981 to 2001. Moreover, the Canadian economy has undergone a significant transformation since the 1980's and it is not surprising to get such results. Ever since the referendum in Quebec, Canadian economy has become more coherent with federal government transferring more and more resources to the provinces. The impact of these programs at the provincial level is more visible in the sample period chosen by our study than in Coulombe and Lee (1995). In short, the increase in convergence rate for this period makes some sense. Moreover, Coulombe and Lee (1995) applied OLS in a pooled regression whereas we have applied a better methodological framework of panel data estimation techniques. This could be another reason for such a divergence in the rate of convergence estimates.

It would, however, be interesting to examine as to what are the driving factors for a faster rate of convergence. Federal equalization program is definitely one of the factors responsible for poorer provinces catching up with the richer ones as some of the studies have already demonstrated (See, for example, Kaufman et al. (2003)). Our intention is to use systems GMM estimator as well to get more robust results. Another important issue may be to divide the sample size used by us into two samples like 1981 to 1990 and 1991 to 2001 and then examine the rate of convergence. One of the most generally accepted results is that the conditional convergence hypothesis holds more strongly when examining more homogeneous group of countries (or regions) (Baumol, 1986, Barro and Salai-i-Martin, 1992, and Mankiw *et al*., 1992 (hereafter MRW)). Using our classification of above average and below averages provinces; we could test the hypothesis of conditional convergence and expect even a higher rate of convergence.

**Acknowledgements:**

We would like to express our sincere thanks to Dr David Giles and Dr. Nilanjana Roy for their invaluable comments on the paper. The usual disclaimer in regard to mistakes applies and these remain the sole responsibility of the authors.

**Footnotes:**

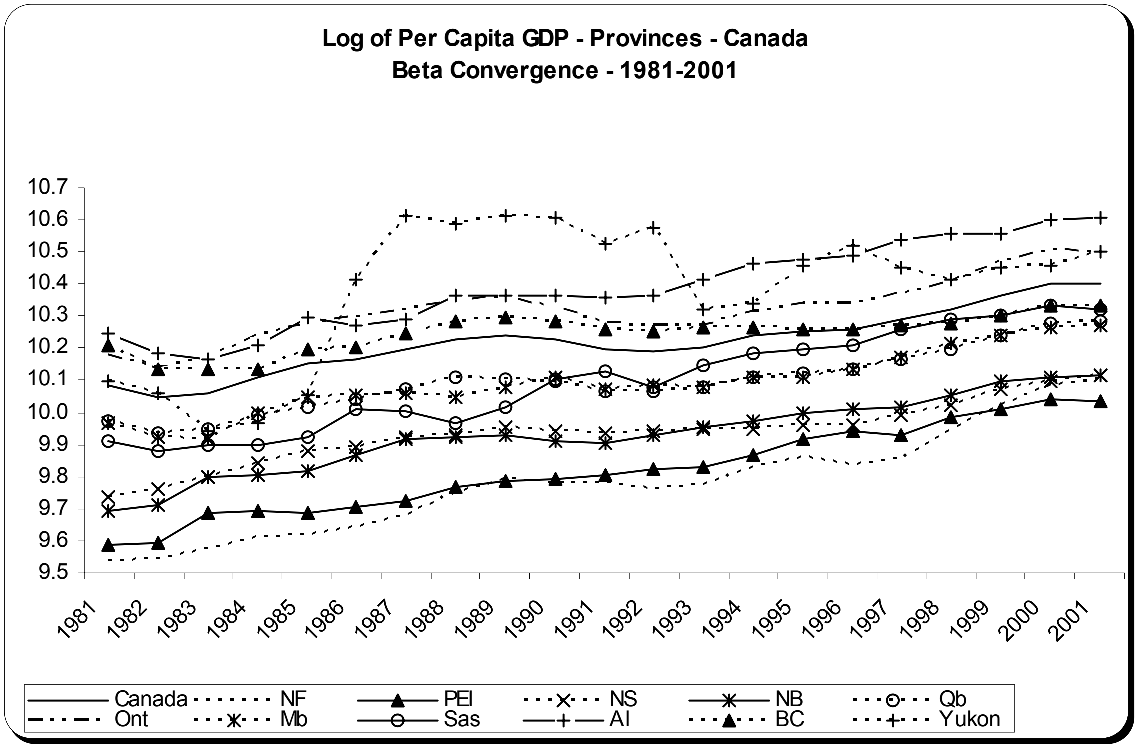
[1] Given $S_t = \rho_1 S_{t-1} + \rho_2 S_{t-2} + ....\rho_{k+1} S_{t-k-1} + \varepsilon_t$, the approximate measure for the time needed to eliminate half of any shock to $\varepsilon$ is $-\ln(2)/\ln(\rho_1)$. The approximate measure is always a real number for our discrete models and so the half-life must be the rounded-up value. To compute the exact half-life, note that any *k+1*[th] order difference equation can be written as *1*[st] order *k+1*[th] vector difference equation of the form $S_t = As_{t-1} + e_t$, where $S_t = \{S_t, S_{t-1}, .....S_{t-k-1}\}'$ and the matrix *A* and $e_t$ are defined accordingly. Then, $ES_T = A_T s_1$. Setting $S_0$ to 0 and then allowing $S_1 = \varepsilon_t > 0$, we can determine the value of *T* that makes $ES_T = \varepsilon_t/2$.

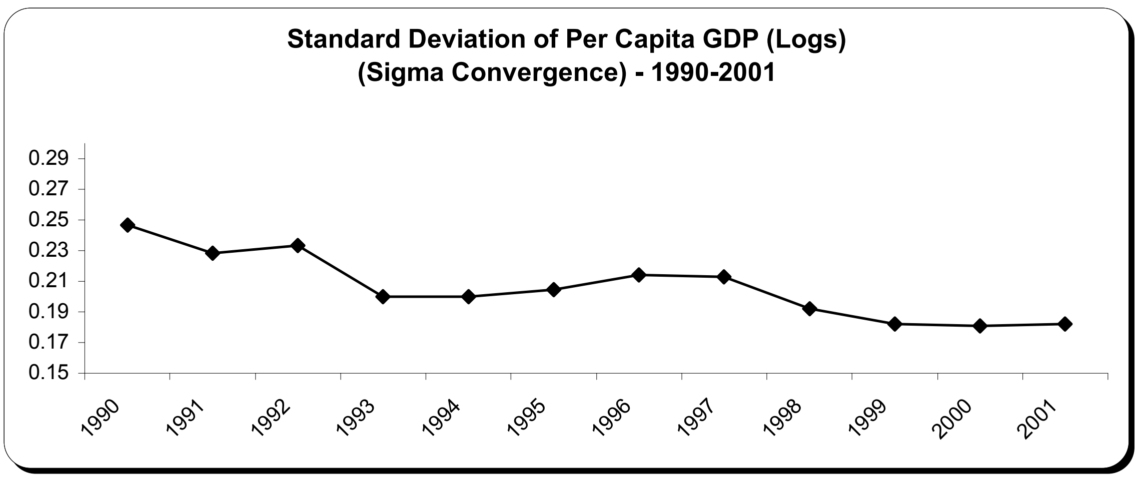
[2] Half-life calculations are based on bias-adjusted $\rho$ estimates, applying Nickel's (1981) formula, which is given by: $p\lim_{n\to\infty}(\hat{\rho} - \rho) = A_T B_T / C_T$; where

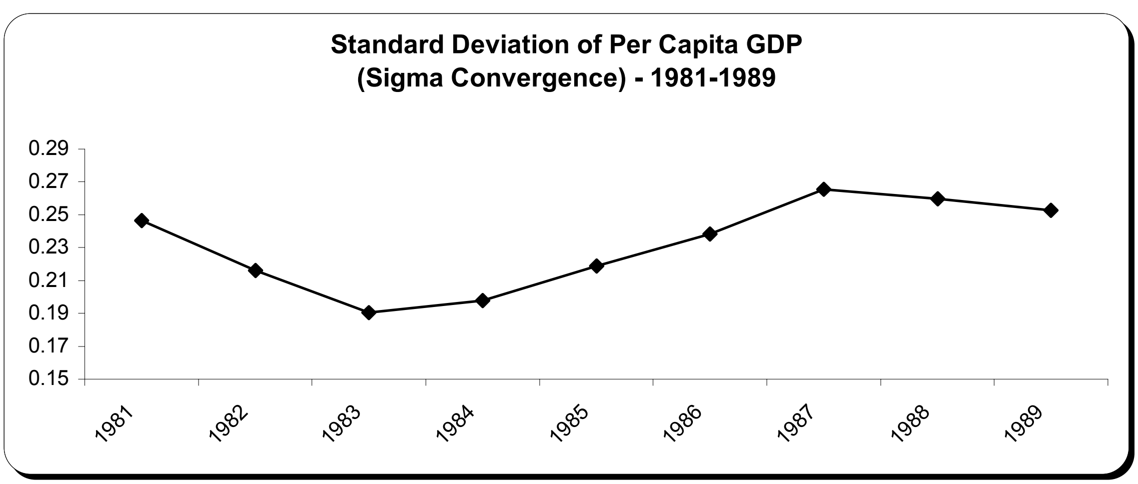
$A_T = -(1+\rho)/(T-1)$, $B_T = 1-(1/T)(1-\rho_T)/1-\rho$),

and $C_T = 1 - 2\rho(1-B_T)/1-B_T)/((1-\rho)(T-1))$ . This bias arises in any AR (1) fixed effects model and is always negative for positive $\rho$. As with the Kendall bias adjustment, we recognize that it is a first order approximation of the bias for an AR (*k+1*) process, and that all *k+1* coefficients will suffer from bias.

## Appendix A


Log of Per Capita GDP - Provinces - Canada
Beta Convergence - 1981-2001
10.7
10.6
10.5
10.4
10.3
10.2
10.1
10.0
9.9
9.8
9.7
9.6
9.5
1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001
Canada
NF
PEI
NS
NB
Qb
Ont
Mb
Sas
Al
BC
Yukon


Standard Deviation of Per Capita GDP (Logs)
(Sigma Convergence) - 1990-2001
0.29
0.27
0.25
0.23
0.21
0.19
0.17
0.15
1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001


Standard Deviation of Per Capita GDP
(Sigma Convergence) - 1981-1989
0.29
0.27
0.25
0.23
0.21
0.19
0.17
0.15
1981 1982 1983 1984 1985 1986 1987 1988 1989

## References:

Arenallo, M. and Bond, S. (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, 58, 277-97.

Barro, R.J. (1991), 'Economic growth in a cross-section of countries', *Quarterly Journal of Economics*, 106, 40743.

Barro, R.J and Salai-i-Martin, X. (1991), 'Convergence across states and regions', *Brookings Papers on Economic Activity*, 1, 107-82.

Barro, R.J and Salai-i-Martin, X. (1992a) Convergence, *Journal of Political Economy*, 100, 223 251.

Barro, R.J and Salai-i-Martin, X. (1992b) Regional Growth and Migration: A Japan-US Comparison, NBER Working paper No. 4308, Cambridge, Mass.

Baumol, W.J. (1986), 'Productivity growth, convergence and welfare: what the long run data show', *American Economic Review*, 76, 1072-85.

Blundell, R. and Bond, S. (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, 87, 115-43.

Coulombe S. and Lee, F.C. (1995), 'Convergence across Canadian provinces, 1961 to 1991', *Canadian Journal of Economics*, 28, 886-898.

Hossain, Akhtar (2000), 'Convergence of per capita output levels across regions of Bangladesh, 1982-97, IMF Working Paper No. 121.

Islam, N. (1995), 'Growth empirics: A panel data approach', *Quarterly Journal of Economics*, 110, 1127-1170.

Kendall, M.G. (1954) "Note on bias in the estimation of autocorrelation", *Biometrika,* 51, 403-404.

Lee, F.C. and Coulombe S. (1995), 'Regional productivity convergence in Canada', *Canadian Journal of Regional Science*, 18, 39-56.

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

Nickell, S. (1981) "Biases in dynamic models with fixed effects", *Econometrica, 49,* 1417-1426.

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

Swan, T.W. (1956), 'Economic growth and capital accumulation', *Economic Record*, 32, 34-361.

Wakerley, E.C. (2002), 'Disaggregate dynamics and economic growth in Canada', *Economic Modelling*, 19, 197-219.

Weeks, M. and Yao, Y. (2003), 'Provincial conditional income convergence in China, 1953 1997: A panel data approach', *Econometric Reviews*, 22, 59-77.