行政院國家科學委員會補助專題研究計計畫成果報告

A Mixture Stochastic Frontier Model and its Application on Economic Growth

計畫類別：個別型計畫
計畫編號：NSC 93－2415－H－032－011－
執行期間： 93年8月1日至　94　年　7　月　31　日

計畫主持人：陳怡宜

成果報告類型(依經費核定清單規定繳交)：精簡報告

執行單位：淡江大學經濟系

中　華　民　國　95　年　12　月　20　日

A Mixture Stochastic Frontier Model and its Application on Economic Growth

1. Introduction

Both the theories and the empirical studies on economic growth have been thriving in the past decade. On the empirical side, an important theme of the literature is on testing the different views of the growth convergence hypothesis regarding either the per capita GDP or the level of GDP. The different views of convergence hypotheses includes the absolute convergence –that poor countries grow faster–, the conditional convergence –that countries farther away from their own steady states grow faster– (Barror and Sala-i-Martin 1992, and Mankiw et al. 1992), and the club convergence –that groups of firms converge to the group-wise steady states (e.x., Durlauf and Johnson 1995). Barro and Sala-i-Martin (1995) provides a comprehensive review.

Different econometric models have been used in the literature to test the convergence hypothesis, and the majority of them are in the linear regression framework. One notable exception is the study by Kumar and Russell (2002). In this seminal work, the authors show how the issue of economic growth can be formulated as a frontier-type problem. The authors then conduct a non-parametric data envelopment analysis (DEA) to investigate issues of growth convergence in the frontier-model framework.

This proposed research project takes the view of Kumar and Russell (2002) and formulates the growth model in a frontier-type framework. This research then departs from Kumar and Russell, and choose to use a parametric stochastic frontier approach (SFA), instead of the non-parametric DEA, to study the frontier-type problem. The difference between SFA and DEA is substantial, and one particularly appealing advantage of SFA over DEA in this particular application is that many of the structural parameters can only be uncovered in a parametric SFA framework.

The SFA model used in this research is a new one to the stochastic frontier literature. The model is the so-called *mixture* stochastic frontier model, which is one of the literature's new developments as discussed in Greene (2001).

This research's contributions to the literature are two folds. One is that we show how the issue of economic growth can be formulated in a parametric SFA, and how the hypotheses of conditional convergence and club convergence can be *simultaneously* tested in this unified framework. The results should be of substantial interests to researchers in the field of economic growth. The other contribution is our development of the mixture SFA model, which shall be of great interests to researchers in the field of stochastic frontier models.

2. The Economic Growth as a Stochastic Frontier Model

Economic growth is a process in which a country approaches its steady-state level of per capita GDP (or GDP level) over time. As argued by Kumar and Russell (2002), the steady-state level of per capita GDP can be regarded as a *frontier* below which lies the actual per capita GDP before convergence. A country can fall short of the frontier because, for instances, inadequate financial institutions or inappropriate regulatory intervention.

The distance between the actual and the frontier per capita GDP changes over time. If the distance shrinks, which will be testable in the econometric model, it implies that the country is converging to its steady state level. If the distance increases, the economy diverges from, rather than converges to, the steady state. In addition, depending on whether the speed of the change is a function of initial income and/or initial distance, the hypotheses of absolute convergence and conditional convergence can be formulated and tested.

2.1. The Role of the Mixture Model

A typical convergence model usually (implicitly) assumes that all the countries share the same preferences, tastes, technologies, etc., so that they all converge to the same steady state. Recent evidence (e.x., Durlauf and Johnson 1995) does not find supportive evidence of this hypothesis; instead, the evidence favors the club convergence hypothesis. In a club convergence, countries with similar cultural, geographical, or economic backgrounds form a relatively homogeneous group, and therefore countries are hypothesized to converge to the group-wise steady states. The implication to the econometric model is that different sets of parameter values should be allowed for different groups of countries in the data.

One way to test the club convergence hypothesis is to classify sample countries into different pre-defined groups, and the researcher then run the growth equation on each of the groups separately. The sample separation is, however, ad hoc. In this proposed research, the mixture model does not require researchers to classify samples *ex ante*. Instead, the sample separation is done endogenously in the estimation process. This approach not only avoids the problematic sample separations, but it is also able to retain all the available data in estimating all the parameters.

Mixture models with normal distributions have a long history in the statistics literature; see McLachlan and Peel (2000) for a comprehensive review. On the other hand, the mixture stochastic frontier model, which has non-normal distributions, is relatively new in the literature (Greene 2001).

In general, the mixture model assumes that agents in the data are heterogeneous, but the discriminating characteristics are not directly observable to econometricians. An endogenous sorting mechanism is then built into the model to reveal certain aspects of the heterogeneity in the estimation results. Therefore, the mixture model seems to be well applicable to study the phenomenon of club convergence.

3. The Econometrics Model

3.1. A Single (not Mixture) Stochastic Frontier Model

Here we lay out the basic stochastic frontier growth model which serves as the bases for the mixture model to be explained in the next section. The basic model is

$$y_{it} = f(k_{it}, z_{it}; \beta) + v_{it} - u_{it}, \tag{1}$$

$$v_{it} \sim N(0, \sigma_v^2), \tag{2}$$

$$u_{it} = G_t \cdot u_i = \exp(\gamma(t - \underline{t})) \cdot u_i, \tag{3}$$

$$u_i \sim N^+(\mu_i, \sigma^2), \tag{4}$$

$$\mu_i = \delta_0 + \delta_1 k_{i\underline{t}}. \tag{5}$$

In this model, y and k are the logarithms of per capita GDP and per capita capital, respectively, and z is a vector of other relevant exogenous variables. The function $f(\cdot)$ measures the *long-run* or *steady state* level of per capita GDP. The variable v_{it} represents random deviations from the long-run level of y. For country i at time t, the term $u_{it} \geq 0$ is the distance that separate the country's log per capita GDP from its steady state. The term u_{it} is a non-stochastic function of time effect (G_t) multiplied by a country-specific, non-negative random variable (u_i). The notation $N^+(\cdot)$ indicates a non-negative truncation of the underlying normal distribution.

Equations (3) to (5) deserve further explanations. In the equations, $\underline{t}$ denotes the initial period of the data, and thus $u_{it} = u_i$ when $t = \underline{t}$. Therefore, the countries' initial distances from the steady state are assumed to follow the distribution of $N^+(\mu_i, \sigma^2)$, and the distance of u_{it} changes over time according to $\exp(\gamma(t - \underline{t}))$. Note that the countries' initial distribution of inefficiencies is assumed to be determined by the country's initial (log of) capital to labor ratio ($k_{i\underline{t}}$) at the initial period, as is in (5). Since the marginal effect of $k_{i\underline{t}}$ on $E(u_i)$ has the same sign as δ_1 (Wang 2002), the prior is that $\delta_1 \leq 0$. That is, countries with lower initial capital labor ratio fell farther below the steady state at the beginning period.

By the specification of (3) to (5), when $t \to \infty$, $u_{it} \to 0$ if $\gamma < 0$. Therefore, whether the per capita GDP converges or diverges can be tested by $H_0 : \gamma < 0$ vs. $H_1 : \gamma \geq 0$. We can further test the conditional convergence hypothesis which says that the growth rate of an economy is positively related to the distance that separates it from its own steady state (Barror and Sala-i-Martin 1992, and Mankiw et al. 1992). To see the implication of conditional convergence on this model, we note that the expected rate of convergence of country i at time t is measured by

$$\theta_{it} = -E\left[\frac{\partial u_{it}}{\partial t}\right] = -\gamma \exp(\gamma(t - \underline{t})) \cdot E(u_i). \tag{6}$$

The first term on the right-hand-side of the equation ($-\gamma \exp(\gamma(t - \underline{t}))$) is common to all the countries, but the second term ($E(u_i)$) is country specific. In particular, if $\delta_1 \leq 0$, the second term is negatively related to $k_{i\underline{t}}$, and therefore $\partial\theta_{it}/\partial k_{i\underline{t}} < 0$ (given $\gamma < 0$). Essentially, the conditional convergence is tested by the joint hypothesis of $\gamma \leq 0$ and $\delta_1 \leq 0$.

If we define $\varepsilon_{it} = v_{it} - G_t \cdot u_i$, and $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iT})'$, then the likelihood function of the model can be derived from

$$\begin{aligned} f(\varepsilon_i) &= \int_0^\infty \Pi_t f(\varepsilon_{it} + G_t u_i) \cdot g(u_i) du_i \\ &= \int_0^\infty \frac{1}{(2\pi)^{T/2}\sigma_v^T} \exp\left(-\frac{1}{2\sigma_v^2}\sum_t (\varepsilon_{it} + G_t u_i)^2\right)\left(\frac{\exp\left(-\frac{(u_i-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma_u \Phi(\mu/\sigma_u)}\right) du_i, \end{aligned} \tag{7}$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution. The above expression can be further simplified to obtain a closed form formula of the likelihood function. Finally, we use ϑ for the vector of parameter to be estimated in the above model, thus the likelihood function of country i will be conveniently denoted as $L_i(\vartheta)$.

3.2. The Mixture Model

Estimation of the above model implies that all the countries share the same structural parameters, and that they all converge to the same steady state. The literature on club convergence argues against this view, and asserts that substantial heterogeneity exists among the countries, so that groups of countries with similar backgrounds converge to the group-wise steady states. The heterogeneity may exist in the form of cultural, technology, weather, and geographic differences.

For the growth model of (1) to (5), the club convergence means that the model parameters vary across groups of countries. We use the mixture model to accommodate such a possibility. Using the subscript j ($j = 1, \ldots, J$) to denote groups, the likelihood function of country i in the mixture stochastic frontier model is

$$L_i(\vartheta, \rho) = \sum_{j=1}^{J} L_{ij}(\vartheta_j) \cdot P_{ij}(\rho_j), \qquad 0 \le P_{ij} \le 0, \qquad \sum_{j=1}^{J} P_{ij} = 1, \tag{8}$$

where $L_{ij}(\vartheta_j)$ is the likelihood function of country i for group j to be derived from (7). The P_{ij} is the probability of country i belonging to group j, and it can be specified as

$$P_{ij} = \frac{\exp(\rho_j q_i)}{\sum_{j=1}^{J} \exp(\rho_j q_i)} \tag{9}$$

where q_i is a vector of country-specific variables, chosen to be determinants of the country's group membership. The ρ_j is the coefficient vector of group j. Parameters of the model can be estimated either by maximizing the likelihood function of (8), or by the EM algorithm (Hartley 1978).

4. Data and Model Specification

The data is mainly obtained from World Development Indicators compiled by World Bank, and the sample period is from 1960 to 1987. The dependent variable in the model is the log of real GDP per capita. The main model specification is the follows.

production frontier variables: $\ln K_{it}$, $\ln H_{it}$, $\Delta \ln N_{it}$; D_{Latin}, D_{EAsia}, D_{SSah}, $\ln NR_i$;

inefficiency variables (μ): $\ln(K/L)_{i0}$;

In the above, $\ln K_{it}$ is the log of the capital stock, $\ln H_{it}$ is the log of human capital, $\Delta \ln N_{it}$ is the growth rate of population. The three dummy variables, D_{Latin}, D_{EAsia}, D_{Ssah}, indicate countries in Latin America, East Asia, and Sub-Sahara, respectively. $\ln NR_i$ is the log of the country's natural resources averaged over the sample period. The variable $\ln(K/L)_{i0}$ is the log of capital to labor ratio measured at the beginning of the sample period.

5. Estimation Results

Table 1 Shows the estimation results with the number of groups equal to 2.

Table 1: Estimated Coefficients

	Group 1		Group 2	
	coef.	t-value	coef.	t-value
constant	11.995	37.090	11.885	25.323
$\ln K_{it}$	0.200	24.638	0.037	7.346
$\ln H_{it}$	0.298	23.347	0.048	5.482
$\Delta \ln N_{it}$	1.095	3.468	0.349	1.184
D_{Latin}	-1.019	-28.667	-1.810	-9.609
D_{EAsia}	0.303	6.153	-0.580	-1.076
D_{SSah}	-1.398	-1.751	-3.668	-13.084
$\ln NR_i$	-0.015	-17.343	-0.010	-4.247
$\ln(K/L)_{i0}$	-0.295	-2.757	-0.918	-1.435
constant	6.179	7.051	7.737	1.941
σ_u^2	0.277	3.367	4.060	1.048
σ_v^2	0.006	39.019	0.012	48.937
γ	-0.010	-22.370	-0.005	-13.279
P	0.463	5.870	0.537	6.339

The results show that physical capitals, human capital, and population accumulations haeve significant beneficial effects on the GDP per capita. Countries in Latin America and, to some extent, Sub-Sahara, have lower GDP per capita given other conditions are equal. Interestingly, countries with larger amounts of natural resources are not benefited from the endowment. The negative and significant coefficient of γ indicates that countries do converge over time. Table 2 lists the countries that are classified into the same groups in the model. The classifications seem reasonable.

Table 2: Country Classifications

Group 1	Algeria, Australia, Austria, Belgium, Brazil, Cameroon, Canada, Colombia, Cote d'Ivoire, Cyprus, Denmark, Ecuador, Finland, France, Greece, Iceland, Indonesia, Ireland, Israel, Italy, Japan, Kenya, South Korea, Malaysia, Mauritius, Mexico, Mozambique, Netherlands, Norway, Paraguay, Portugal, Singapore, Spain, Sweden, Thailand, Tunisia, USA.
Group 2	Argentina, Bangladesh, Bolivia, Chile, China, Congo, Costa Rica, Egypt, Ethiopia, Ghana, Guatemala, Haiti, Honduras, India, Iran, Jamaica, Jordan, Madagascar, Malawi, Mali, Morocco, New Zealand, Nigeria, Pakistan, Peru, Philippines, Rwanda, Senegal, Sierra Leone, Sri Lanka, Switzerland, Tanzania, Turkey, Uganda, United Kingdom, Uruguay, Venezuela, Zambia, Zimbabwe.

Remarks

We also estimate models with **P** being parameterized by a list of exogenous variables. The results are similar to the one reported above. On the other hand, when estimating the model with more than three groups, numerical difficulty arises. This may not a serious limitation since many studies in the literature find countries converge to no more than two groups (e.x., Kumar and Russell 2002).

References

[1] Barro, Robert, and Sala-i-Martin, Xavier. (1992). Convergence, *Journal of Political Economy* 100, pp.223-51.

[2] Barro, Robert, and Sala-i-Martin, Xavier. (1995). Economic growth. Advanced Series in Economics. New York; London and Montreal: McGraw-Hill.

[3] Durlauf, Steven N., and Johnson, Paul A. (1995). Multiple Regimes and Cross-Country Growth Behaviour, *Journal of Applied Econometrics* 10, pp.365-84.

[4] Greene, William. (2001). New Developments in the Estimation of Stochastic Frontier Models with Panel Data, Stern School of Business, New York University.

[5] Hartley, M. (1978). Comment on Estimating Mixtures of Normal Distributions and

Switching Regressions, *Journal of American Statistical Association* 73, pp.738-741.

[6] Heston, Alan, Summers, Robert, and Aten, Bettina. (2002). Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP).

[7] Kumar, Subodh, and Russell, Robert. (2002). Technological Change, Technological Catch-Up, and Capital Deepening: Relative Contributions to Growth and Convergence, *American Economic Review* 92, pp.527-48.

[8] Mankiw, N. Gregory, Romer, David, and Weil, David N. (1992). A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107, pp.407-37.

[9] McLachlan, Geoffreg. (2000). Finite Mixture Models. New York: John Wiley & Sons.

[10] Wang, Hung Jen. (2002). Heteroscedasticity and Non-monotonic Efficiency Effects of a Stochastic Frontier Model, *Journal-of-Productivity-Analysis* 18, pp.241-253.