

A NONPARAMETRIC ANALYSIS OF THE PERSONAL INCOME DISTRIBUTION ACROSS THE PROVINCES AND STATES IN THE U.S. AND CANADA

WANG, Yibing*

Abstract

A nonparametric stochastic kernel and Markov transition probability matrixes are used in this paper to investigate the per capita personal income distribution and how it evolved over the period 1950 - 2000 across the 59 provinces and states in the U.S. and Canada. Empirical evidence confirms club convergence (multi-modality) of the per capita personal income levels across the provinces and states in the U.S. and Canada over the entire study period.

JEL classification: C1, D31, O51

Key words: Convergence, stochastic Kernel and Markov chains

1. Introduction

In recent years there has been an increasing literature on income convergence analysis that makes use of nonparametric methods. Convergence implies a long-run tendency towards the equalization of *per capita* income levels. In other words, convergence addresses the important question of whether the per capita income levels in poor economies (countries or regions) will catch up with those in rich economies.

By estimating non-parametrically the cross-country income data, Bianchi (1997) and Quah (1997) tested the growth convergence and

* Yibing Wang is Assistant Professor., Economics Department, Shandong University at Weihai, Shandong China. Email: wang_yibing@yahoo.com

Acknowledgement: I would like to thank to the Editor, Maria Carmen Guisan, for her very helpful suggestions and comments that led to substantial improvements of this paper.

*convergence clubs*¹ hypothesis. They found evidence of “twin peaks” (*bimodality*) in the world relative income distribution via a thorough analysis of the changes in the shape of the income distribution. Applying a tree-clustering algorithm on cross-country data, Durlauf and Johnson (1995) divide the countries into four different groups with different growth behavior. Multiple steady states (*multi-modality*) in cross-country growth behavior were found in their regression results. From the above one may realize that the test of multi-modality can be very useful in the research of income distributions. Important hypotheses such as income convergence can be formulated as a question of multi-modality.

The objective of this paper is to investigate the convergence (unimodality) issues of the personal income distribution across the 59 provinces and states in the U.S. and Canada by applying two nonparametric approaches, specifically, stochastic kernels and Markov transition probability matrixes.

The rest of the paper is outlined as follows. Section II provides a review of the conventional parametric estimation methods of the convergence of income distribution and a reasonable critique of these classic methods. Section III gives a detailed description of the kernel methodology used in this study. Section IV briefly discusses the source and quality of our income data. In section V stochastic kernel and Markov transition probability matrix are used to analyze the income distribution dynamics. Section VI concludes and gives future research suggestions.

2. Classical convergence concepts and critics

Numerous attempts have been made to provide a precise description of the convergence of income. There are three major classical

¹ Baumol (1986) and Quah (1993) have all been strong proponents of the idea of grouping similar countries or economies into different *convergence clubs*, such as the rich country club, the poor country club etc. While overall convergence may not exist, countries within a convergence club may show signs of convergence (*club convergence*).

convergence concepts: 1) Absolute *b*-convergence 2) Conditional *b*-convergence 3) Sigma (σ)-convergence.

1) The absolute *b*-Convergence

The *absolute b-Convergence* approach, developed mainly by Baumol (1986) and Barro and Sala-i-Martin (1991, 1992), is to estimate cross-section growth rates on initial levels of income as follows:

$$\log(y_{i,t+T} / y_{i,t}) / T = a + b \log(y_{i,t}) + e_{i,t} \quad (1)$$

where $\log(y_{i,t+T} / y_{i,t}) / T$ is economy *i*'s growth rate of per capita income between *t* and *t+T*, *T* is the length of time over which the growth of per capita income is measured, and $e_{i,t}$ is a stochastic error term, *a* is a constant term representing the steady-state point of convergence which is the same for all economies and *b* is the convergence coefficient. According to the neo-classicists, a negative sign of the *b* coefficient indicates that the growth rates in per capita incomes over the *T* year period were negatively correlated with starting incomes, or in other words the initially poor regions grow faster than the initially richer ones.

Absolute *b-convergence* has been widely tested for the U.S. regional convergence. Sala-i-Martin (1991, 1992) found evidence in favor of absolute *b-convergence* in both the per capita personal income and per capita gross state product with an estimated convergence speed at about 2% per annum for the period 1963-1986.

Sala-i-Martin (1996) pointed out that the convergence of all economies to the same steady state predicted by the neoclassical model relies heavily on the assumptions that the only difference across countries lies in their initial levels of capital. In reality, however, economies may differ in levels of technology, propensities to save, or population growth rates. With these differences in levels of technology and preferences, different economies most likely will have different steady states and the absolute *b-convergence* will be flawed by imposing

the restriction that they are the same.

2) Conditional *b*-Convergence

To overcome the obvious flaw in the absolute *b*-estimation of the classical growth theory, authors such as Barro and Sala-i-Martin (1992), Mankiw, Romer and Weil (1992) and Durlauf (1996) among others, developed the idea of conditional *b-convergence*.

Conditional *b*-convergence argues that economies converge to different steady-state points of growth since they have different economic structures. Convergence is conditional on the steady-state growth path that is a function of the differences in technology levels, human capital, investment and saving rates, among other structural variables.

To test the hypothesis of conditional *b*-convergence one has to extend equation (1) of the absolute *b*-convergence to incorporate some structural variables, such as the investment ratio, human capital, innovative activity, public expenditure, population growth, trade, and so on. In other words, instead of estimating (1) one estimates

$$\log(y_{i,t+T} / y_{i,t}) / T = a + b \log(y_{i,t}) + yX_{i,t} + e_{i,t+T} \quad (2)$$

In equation (2), all the variables are the same as in equation (1), and $X_{i,t}$ is a vector of structural variables (as proxies for the steady state) that affect the steady state of economy *i*. If the estimate of *b* is negative once $X_{i,t}$ is held constant, then the data set is said to exhibit conditional *b* -convergence.

Conditional *b*-convergence has also been extensively tested for the U.S. regional convergence. Sala-i-Martin (1992) included sector shift parameters as well as regional dummies to hold constant possible shocks and found evidence of conditional *b*-convergence at a rate of about 2% per annum in per capita gross state product for the period 1963-1986.

A number of econometric problems have been identified with conditional *b*-convergence analysis. The initial level of technology, which should be included in a conditional *b*-convergence specification, is not observed. Since it is also correlated with other regressors (such as initial income), the conditional *b*-convergence studies suffer from an omitted variable bias.

3) *Sigma (s)-Convergence*

The concept of *sigma-convergence* is primarily developed by Baumol (1986), Dowrick and Nguyen (1989), Barro and Sala-i-Martin (1991, 1992). The central idea is that if the standard deviations of per capita incomes across groups of economies exhibit a decreasing trend over time, then there is sigma-convergence. Formally, following Sala-i-Martin (1996), sigma-convergence occurs if

$$s_{t+T} < s_t, \quad (3)$$

where s_t is the time t standard deviation of $\log(y_{i,t})$ across i and T is the time period over which the standard deviation is observed.

The *s*-convergence has been extensively tested for the U.S. case. Mitchener and McLean (1999) studied the US state personal income for 1880-1980 and found different sigma-convergence rates with different choices of series (nominal income, price adjusted income or labor productivity).

As argued by Quah (1996), the main flaw of *s*-convergence is that when s_t is constant over a study period, thus signaling no convergence or divergence, the underlying economies may actually still be moving within an invariant distribution frame.

The above classic convergence concepts are frequently used to analyze the regional income distributions, which involve imposing parameter functions. Unfortunately, all the classic parameter functions explained so far are with defects in some ways.

3. Non-parametric density estimation

Unlike the parametric approach, nonparametric density estimation allows one to draw a complete picture and hence provides full information on the entire income distribution.

Nonparametric estimation approach allows one to analyze data at hand without any *a priori* assumptions on the form of the underlying density of the data. The only requirement about the data, if any, is perhaps that the underlying density of the data should be smooth enough for meaningful analysis.

Although various parametric methods have been extensively used to analyze the world income distribution and growth convergence. An in-depth analysis of the per capita personal income across provinces (states) in *both* the U.S. *and* Canada based on nonparametric stochastic kernels has not yet been found in the previous studies on income distribution dynamics.

3.1 Kernel Estimator

The kernel estimator is one of the most popular nonparametric estimators due to its simplicity to calculate and interpret.

Let $f = f(x)$ denote the continuous density function of a random variable X at a point x , and let $x_1, \dots, x_n$ be the observations from f .

Rosenblatt (1956) defined a kernel function K as:

$$\int_{-\infty}^{\infty} K(y) dy = 1 \quad (4)$$

where $K(y) \geq 0$.

The general kernel estimator $\hat{f}(x)$ of $f(x)$ is defined by:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) = \frac{1}{nh} \sum_{i=1}^n K(y_i), \quad (5)$$

where $y_i = h^{-1}(x_i - x)$, n is the number of observations in the sample,

h is the window-width (bandwidth) which is a function of the sample size and goes to zero as $n \rightarrow \infty$.

3.2 Choice of Bandwidth h

Hardle (1990), Silverman (1986) and others argued that the choice of kernel is not crucial to analysis since any kernel could be optimal for large enough samples. In contrast the selection of the window width h is critical. There is always a trade-off between bias and variance when choosing window width h . On the one hand, a very small h may result in under-smoothing that may reduce the bias of the estimator but increase its variance. A large h , on the other hand, could result in over-smoothing that reduces variances but increases bias, besides missing important details about the distribution. Hence great caution should be executed in selecting the window width h .

One has to have some criterion in order to make the optimal choice of h . So far the most popular method (originally introduced by Rosenblatt (1956)) has been to minimize $E \left\{ \int [\hat{f}(x) - f(x)]^2 dx \right\}$, the mean integrated squared error (*MISE*).

In practice, $AMISE^2$ (Asymptotic *MISE*) is used to approximate *MISE* which is difficult to calculate.

$$\begin{aligned} MISE(\hat{f}) &\approx AMISE(\hat{f}) = \frac{h^4}{4} m_2^2 \int (f^{(2)}(x))^2 dx + (nh)^{-1} \int f(x) dx \int K^2(y) dy \\ &= \frac{h^4}{4} m_2^2 \int (f^{(2)}(x))^2 dx + (nh)^{-1} \int K^2(y) dy \end{aligned}$$

Silverman (1986) showed that the value of h that minimizes *AMISE* is

$$h = 0.9 [\min(\hat{s}, R/1.34)] n^{-1/5},$$

² Proof of this can be found in Pagan and Ullah (1999): pp 22-24

where R stands for the inter-quartile range and $\hat{s}$ is the sample standard deviation.

4. Data: sources, transformation and sub-periods division

Fifty states in the U.S. and nine provinces in Canada are included in this study³. The per-capita personal income over the period 1950 - 2000 for the 59 regions in the U.S. and Canada is used to analyze the regional income distribution in this paper.

The data of the regional per capita personal income in the U.S. is taken from the well-known and widely used Bureau of Economic Analysis Regional Accounts Data (1929 – 2001) and the data of the provincial per-capita personal income in Canada is extracted from CANSIM (Canadian Socio-Economic Information Management System)⁴. CANSIM is the Statistics Canada's computerized database of time series covering a wide variety of social and economic aspects of Canadian life.

The per-capita personal income in each year for each region is normalized by the average per-capita personal income for the total 59 economies in that year (with the average taking a value of 1.00).

Since one of the purposes of this study is to compare per capita personal income across years, incomes are reported in 1982-84 prices by correcting for inflation using the consumption price index (CPI)⁵.

In order to make the provincial personal income in Canada comparable to that in the U.S., the personal incomes in Canada are converted to U.S. dollars as per average exchange rate in each year.

³ See appendix TableA.1 (unadjusted for inflation) for the provinces and states.

⁴ CANSIM is Statistics Canada's computerized database and information retrieval service. Data before 1990 are from CANSIM II SERIES V501122, data after 1990 are from CANSIM II SERIES v691825.

⁵ The Canada CPI and the U.S. CPI are from CANSIM Series P10000 and D19805 respectively.

5. A Preliminary look at the income distribution

Figure 1 presents the entire distribution of per capita personal income (all relative to the North American average, excluding Mexico) across 9 Canadian provinces and 50 U.S. states for the period 1950 – 2000.

The North American average is indicated by one on the vertical (Z) axis marked *Relative Income* in Figure 1. Time periods are sequentially marked along the X-axis of *Year*. Different provinces and regions are represented along the Y-axis marked *States and Provinces*.⁶

The Canadian provinces are indicated by vertical lines along the Y-axis at 1.0 (Prince Edward Island), 2.0 (New Brunswick), 3.0 (Nova Scotia), 4.0 (Saskatchewan), 5.0 (Quebec), 8.0 (Alberta), 9.0 (Manitoba), 14.0 (Ontario) and 15.0 (British Columbia) respectively. It can be noted that in 1950 the relative incomes in all the Canadian provinces are below the North American average.

From Figure 1, it is clear that the pure use of either cross-sectional distribution or time-series distribution would fail to provide complete intra-distribution information. From Figure 2, one may notice that the income levels in the Canadian provinces experienced a slow convergence to the North American average from 1950 to 1970, but became approximately stagnated between 1970 and 1990, then diverged from the North American average over the 1990's. It is also interesting to note that Ontario, British Columbia and Alberta have almost always been the three highest income provinces over this whole study period and all experienced income levels higher than the North American average for some years between 1970 and 1990.

6. Income distribution dynamics

Figure 3 tracks the evolution of the distribution of the relative per capita personal income across the 59 US-Canada regions through the

⁶ Provinces and states are ordered from the lowest per capita income to the highest per capita income in 1950 as shown in the appendix Table A1.

period 1950 – 2000. The middle line shows the 50th percentile (median) of the relative income distribution in each year. The top and bottom lines show, respectively the 75th and 25th percentiles of the

Figure 1. Relative incomes across 59 Provinces and States in the U.S. and Canada for 1950-2000

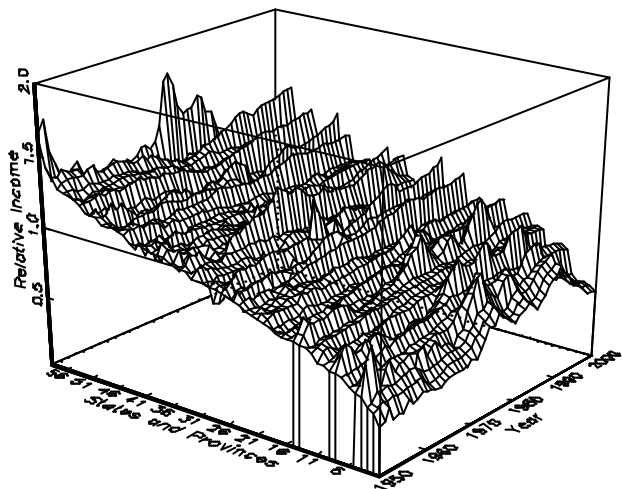
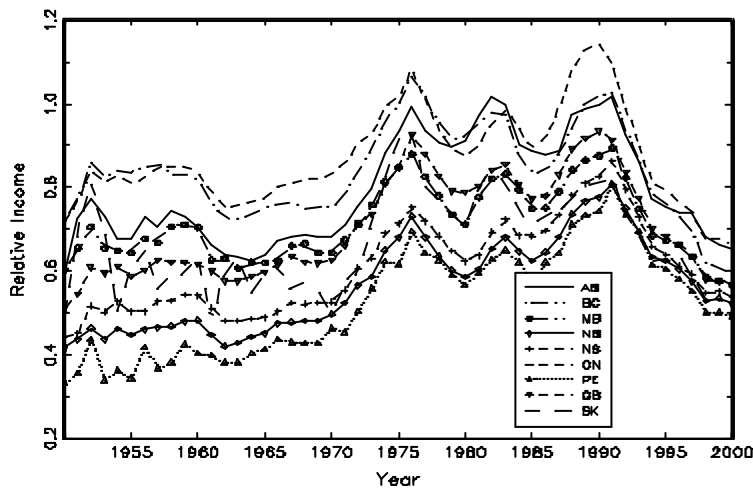


Figure 2. Relative incomes in Canadian provinces for 1950 -2000



relative income distribution. The distance between 75th percentile line and the 25th percentile line indicates the inter-quartile range (*IQR*).

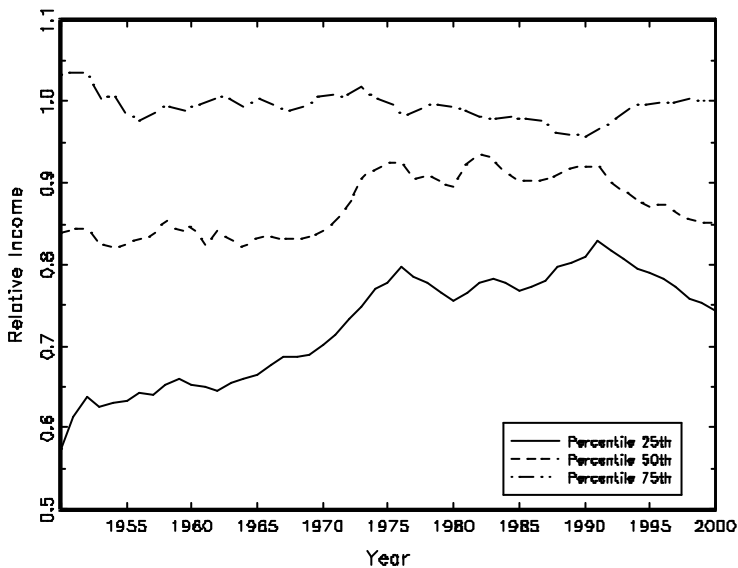
The *IQR* is defined as

$$IQR = Q_3 - Q_1$$

where Q_3 indicates the 75th percentile and Q_1 represents the 25th percentile.

From 1950 to 1970, the 25th percentile group experienced a significant increase in its relative income while the 75th percentile and the 50th percentile remained approximately unchanged. Over the following period 1971-1990, the relative incomes of the 75th percentile group and the 50th percentile group were fairly stable without significant change while the relative income of the 25th percentile group became stable after a sharp increase in the beginning of this period. In the 1990's, the 25th and the 50th percentile group experienced some decreases and the 75th percentile group was rising.

Figure 3. Evolution of Relative Income



Another fact derived from Figure 3 is that the cross-sectional distribution of income is quite tight, centered on approximately 0.9. The distance of the lines from each other also provides important information about the income gap. The closer are the lines, the lower is the income gap. Figure 3 suggests personal income gap in the U.S. and Canada experienced a period of rapid decrease from the 1950's to the 1970's, but slow increase in the 1990's.

6.1. Three-Dimensional Representations: The Stochastic Kernel⁷

Quah (1996) argued that a stochastic kernel (as well as its contour) is a graphical representation of the transition probabilities with the advantage that it gives estimates for continuous states of transition probabilities.

Let $\{x_t, x_n\}$ be a set of income data at time t from density $f_t(x)$, and after period k the income changes to $\{y_t, y_n\}$ and its corresponding density evolves to $f_{t+k}(y)$. Then the relation between these two densities can be described by:

$$f_{t+k}(y) = \int_0^{\infty} T_k(y|x) f_t(x) dx \quad (6)$$

where $T_k(y|x)$ is the stochastic kernel (transition probability) that could be used to describe the income distribution from time t to time $t+k$.

Following Quah (1997), a stochastic kernel conveys important information on income distributions. In this paper, the relative incomes in period t and $t+k$ are respectively represented along the t axis and the $t+k$ axis. The 45-degree diagonal line represents constant income. Thus, points lying along the 45-degree line indicate that relative incomes remain unchanged, while points to the left (right) of the diagonal signal a rise (decrease) in relative incomes between any two periods studied. *Unimodality* in a stochastic kernel indicates convergence while *multi-modality* may indicate club convergence.

⁷ The stochastic kernel is a conditional density function. Estimation of the kernel is carried out by first estimating the joint density function of the process at time t and $t+k$ and then normalizing it by the marginal in t . See Quah (1996a).

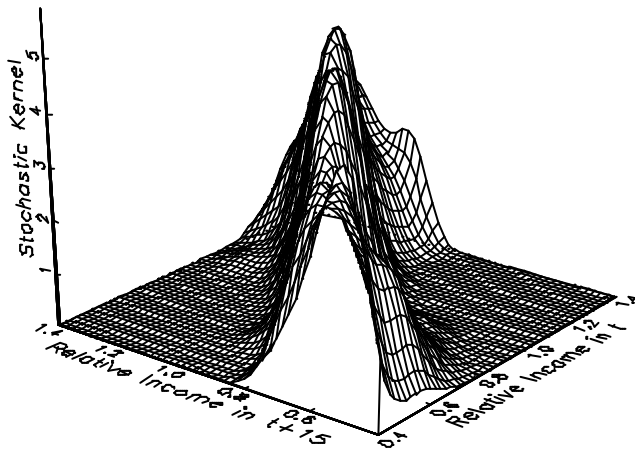
Furthermore, the tendency to convergence is also indicated if most of the stochastic kernel mass becomes more parallel to the t axis by making a clock-wise movement around the center of the 45-degree line while the tendency to divergence is implied if most of the stochastic kernel mass becomes more vertical to the $t + k$ axis by making a counter-clockwise movement around the center of the 45-degree line.

6.2. Empirical Evidence from the Stochastic Kernel

Figure 4 ~ Figure 5 present the stochastic kernel and contour plots for the 59 regions in the entire study period (1950 – 2000)⁸.

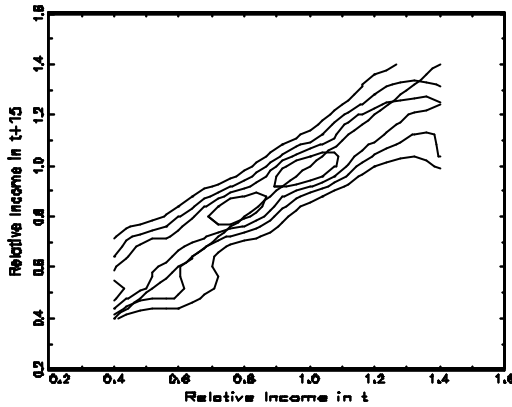
There are two distinct features in the stochastic kernel of the entire period: twin peaks (bimodality) – the average income peak at 1.0 and the low-middle income peak at about 0.8. Since most mass of the stochastic kernel is on the 45-degree line, one can conclude mobility of income distributions is not very strong throughout the whole study period. It seems regions tend to converge to different steady states and exhibit a form of club-convergence in the long run.

Figure 4. Stochastic Kernel 1950 - 2000



⁸ The stochastic kernels are of transitions of 6 years, 6 years, 3 years and 15 years respectively for each study period. Transition of years can be set at other numbers and will give similar results.

Figure 5. Contour for Figure 14



6.3. Markov Transition Probability Matrix (MTPM)

Another traditional nonparametric way to describe income distribution dynamics is transition probability matrix, which is actually the discrete version of the stochastic kernel. A simple first order Markov chain can be described as

$$X_{t+1} = M_{t+1} X_t$$

where X_t is a vector giving the number of economies in each income state at time t and X_{t+1} is a vector a period later. If the probabilities of the economies moving between different income states from period t to $t+1$ can be described by M_{t+1} , then M_{t+1} is a transition probability matrix. In practice, a transition probability matrix is estimated by using discrete states of income distributions. Following Quah (1993), the first column in a transition probability matrix table gives the total numbers of observations in a specific state in a study period, the first row provides the upper points of the corresponding cells. Each row i gives the estimated probabilities of staying in that state i and of moving from state i to other states. The last row presents the ergodic distribution. Numbers along the diagonal indicate the level of immobility while off-diagonal numbers imply the degree of mobility.

6.4 Results from MTPM

Table 1 presents one-year horizon transition probabilities between different income states for the period 1950 - 2000. The values in the main diagonal of Table 1 for the one-year transition are around 90%, indicating a high degree of immobility during this period of study. This finding is consistent with the previous results: the stochastic kernel for the period 1950 – 2000 indeed exhibits a low degree of mobility.

Mobility is considerably higher in the second state than in other states of the income distribution. It shows that a region in state 2 has 10% probability to move to the lower state and 11% probability to move to the higher state 3. This indicates that there is a strong tendency for the upper-low-middle income regions to move out of their initial income states. The ergodic distribution was reported in the last row of the table. The ergodic distribution indicates the long run tendency of an economy staying in a given state regardless of its initial state.

The results indicate that over the long run, the probability of an economy staying in the 3^d state is the highest, a little over 32% and the probability of landing in the 1st state is the second highest, nearly 30%. This is encouraging as it indicates that overall the middle-income regions are stable in the long term in the U.S and Canada despite a form of club convergence. Similar results are presented in Table 2, which gives a five-year horizon transition probability between different income states for the same period. It shows that the degree of mobility is higher over a five-year transition period.

The results for the second state group are still noticeable in that they show a probability of a region moving out of its own initial state is much higher than that of any other regions. In sum, the results from Markov transition probability confirm the previous findings in section 6.2.

Table 1. Markov Transition Probability Matrix, 1-Year Horizon (1950 – 2000)

Number	Upper endpoint			
	0.85	0.95	1.1	1.6
(931)	0.94	0.06		
(536)	0.10	0.79	0.11	
(855)		0.07	0.90	0.03
(563)			0.06	0.94
Ergodic	0.292	0.191	0.325	0.193

Table 2. Markov Transition Probability Matrix, 5-Year Horizon (1950 – 2000)

Number	Upper endpoint			
	0.85	0.95	1.1	1.6
(873)	0.83	0.15	0.02	
(536)	0.17	0.57	0.25	
(855)	0.02	0.12	0.77	0.09
(563)	0.01		0.14	0.85
Ergodic	0.233	0.185	0.356	0.226

7. Conclusions and suggestions for future research

Both stochastic kernels and Markov transition probability matrixes are applied to examine the long run state (provincial) per capita personal income distribution dynamics, such as convergence and mobility, across the provinces and states in the U.S. and Canada over the period 1950-2000. The entire study period shows a form of club convergence.

Income gap experienced a rapid decrease in the 1950-1970 period and kept almost unchanged in the 1971-1990 period. An obvious reversal to this general trend is in 1991-2000, a decade marked with remarkable economic expansion and the advent of a “New Economy”. This paper shows the merits of nonparametric methods in analyzing income distributions when multi-modality is present. Most federal governments make policies to eliminate regional income disparities on the basis of results from classical parametric methods, but this may be

inappropriate because, for example, a mean value (first moment), such as of income, from a bimodality density distribution does not provide the same inferences as that from a unimodality density distribution.

One limitation of this paper is that it does not provide economic rationales as why bimodality may or may not occur in different periods. The goal of further research will be to apply the nonparametric methods employed in this paper to analyze the per capita personal income distribution conditional on variables (such as saving rates, population growth rates, trade openness, physical neighbors and taxation levels etc.) that can provide additional information on income distributions.

Appendix: Table A1. Provinces and States in Canada and the U.S.

Region	1950(US\$)	2000(US\$)	Region	1950 (US\$)	2000 (US\$)
P.E. Island	533.6(59)	14642.8(59)	North Dakota	1360.0(30)	25007.0(38)
N. Brunswick	673.7(58)	15576.1(57)	Texas	1363.0(29)	28035.0(24)
Nova Scotia	714.8(57)	16017.1(56)	Arizona	1367.0(28)	25358.0(37)
Saskatchewan	806.4(56)	15503.3(58)	Missouri	1427.0(27)	27452.0(29)
Quebec	824.7(55)	16879.0(54)	Hawaii	1429.0(26)	28301.0(20)
Mississippi	770.0(54)	21017.0(50)	Minnesota	1437.0(25)	32207.0(10)
Arkansas	847.0(53)	22108.0(47)	Kansas	1463.0(24)	27537.0(28)
South Carolina	925.0(52)	24273.0(39)	Wisconsin	1506.0(23)	28471.0(19)
Alabama	909.0(51)	23766.0(43)	Colorado	1521.0(22)	33018.0(7)
Alberta	967.5(50)	19505.8(52)	Indiana	1524.0(21)	27228.0(31)
Manitoba	975.7(49)	16872.9(55)	Iowa	1532.0(20)	26572.0(33)
Kentucky	990.0(48)	24244.0(40)	Pennsylvania	1552.0(19)	29713.0(15)
Tennessee	1028.0(47)	26367.0(34)	Rhode Island	1553.0(18)	29258.0(18)
Ontario	1163.3(46)	19752.2(51)	Nebraska	1560.0(17)	27756.0(27)
B.C.	1173.4(45)	17609.6(53)	Ohio	1608.0(16)	28202.0(23)
West Virginia	1056.0(44)	21861.0(48)	Maryland	1642.0(15)	33959.0(5)
Georgia	1065.0(43)	28212.0(22)	Montana	1654.0(14)	22895.0(46)
North Carolina	1077.0(42)	27055.0(32)	Massachusetts	1656.0(13)	37960.0(2)
Louisiana	1117.0(41)	23227.0(45)	Oregon	1657.0(12)	27821.0(25)
Oklahoma	1144.0(40)	24046.0(42)	Michigan	1718.0(11)	29516.0(17)
Vermont	1169.0(39)	27376.0(30)	Wyoming	1719.0(10)	27767.00(26)
Maine	1195.0(38)	25681.0(36)	Washington	1721.0(9)	31627.0(11)
New Mexico	1204.0(37)	21837.0(49)	New Jersey	1802.0(8)	37649.0(3)

Virginia	1257.0(36)	31320.0(13)	Illinois	1831.0(7)	32248.0(9)
South Dakota	1283.0(35)	25823.0(35)	New York	1858.0(6)	35016.0(4)
Florida	1304.0(34)	28286.0(21)	California	1877.0(5)	32334.0(8)
Idaho	1329.0(33)	24101.0(41)	Connecticut	1891.0(4)	41392.0(1)
N. Hampshire	1348.0(32)	33576.0(6)	Nevada	1991.0(3)	29696.0(16)
Utah	1348.0(31)	23476.0(44)	Delaware	2075.0(2)	31500.0(12)
			Alaska	2400.0(1)	29913.0(14)

Notes 1) Numbers inside brackets are ranks of per capita income levels. 2) The population growth rates of the provinces and states are relatively stable and no structural break is present in the population growth, therefore the effect of changes in population growth on convergence is neglected. 3) Canadian provinces are in bold letters. Canadian province Newfoundland is not included in this paper as it has long been isolated from the main Canada economy and its data may suffer from major structure changes and are unreliable.

References

Barro, R. (1991), "Economic Growth in a Cross Section of Countries", *The Quarterly Journal of Economics*, 106 (2): 407-443

Barro, R. and Sala-I-Martin, X. (1991), "Convergence Across States And Regions", *Brookings Papers on Economics Activity*, 1991 (1): 107-182.

Barro, R. and Sala-I-Martin, X. (1992), "Convergence", *Journal of Political Economy*, 100 (2): 223-251

Barro, R. and Sala-I-Martin, X. (1995), *Economic Growth*. New York: McGraw-Hill, 1995

Baumol, W. (1986), "Productivity Growth, Convergence and Welfare. What Long-Run Data Show", *American Economic Review*, 76(5): 1072-85, December.

Bernard, A. B., and Durlauf, S.N (1995), "Convergence in international output", *Journal of Applied Econometrics*, 10 (2): 97-108

Bernard, A. B., and Durlauf S.N. (1996), "Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, 71 (1-2): 161-174

Bianchi, M. (1997), "Testing for Convergence: Evidence from

Nonparametric Multimodality Tests,” *Journal of Applied Econometrics*, Vol. 12, pp. 393-409.

Dowrick, S. and Nguyen, D. (1989), “OECD comparative economic growth 1950-85: catch-up and convergence”, *American Economic Review*, 79 (5): 1010-30

Durlauf, S.N. (1989), “Output persistence, economic structure, and the choice of stabilization policy”, *Brookings Papers on Economic Activity*, 1989 (2): 69-116

Durlauf, S.N. (1996), “A Theory of persistent income inequality”, *Journal of Economic Growth*, 1 (1): 75-93

Durlauf, S. and JOHNSON, P. (1995), “Multiple Regimes and Cross-Country Growth Behaviour,” *Journal of Applied Econometrics*, Vol. 10, pp. 365-384.

Durlauf, S. and QUAH, D. (1999), “The New Empirics of Economic Growth,” *Handbook of Macroeconomics*, Vol. 1A, Amsterdam: Elsevier Science.

Hardle, W. (1990), *Applied Nonparametric Regression*, New York, Cambridge University Press

Jones, C. (1997), “On the Evolution of the World Income Distribution,” *Journal of Economic Perspectives*, Vol. 11:3, pp. 19-36.

Mankiw, N., Romer D. and. Weil D. (1992), “A contribution to the empirics of economic growth”, *Quarterly Journal of Economics*, 107(2): 407-437

Mitchener, K. and Mclean W. (1999), “U.S. Regional Growth And convergence, ”1880-1980, *The Journal of Economic History*

Pagan, A. and Ullah, A. (1999), *Nonparametric Econometrics*, Cambridge: Cambridge University Press.

Quah, D.T. (1993), "Empirical cross-section dynamics in economic growth", *European Economic Review*, 37(2/3): 426-434

Quah, D. (1996a), "Convergence Empirics Across Economies with (Some) Capital Mobility", *Journal of Economic Growth*, 1 (1):95-124

Quah, D. (1996), "Empirics for economic growth and convergence", *European Economic Review*, 40 (6): 1353-1375

Quah, D. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, Vol. 2, pp. 27-59.

Rosenblatt, M (1956), "Remarks on Some Nonparametric Estimates of a Density Function", *Annals of Mathematical Statistics*, 27, 832-837

Sala-I-Martin, X. (1994). "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence," *Economic Growth Center Discussion Paper*, 715:46

Sala-I-Martin, X.(1996), "The Classical Approach to Convergence Analysis," *The Economic Journal*, 106, 1019-1036

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, New York, Chapman and Hall