CONVERGENCE ACROSS PROVINCES OF TURKEY: A SPATIAL ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ALTAN ALDAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
DEPARTMENT OF ECONOMICS

JULY 2005

Approval of the Graduate School of Social Sciences

Prof. Dr. Sencer AYATA
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Erol Çakmak
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Esma GAYGISIZ
Supervisor

Examining Committee:

Prof. Dr. Erol TAYMAZ (METU, ECON) ____________________

Assist. Prof. Dr. Esma GAYGISIZ (METU, ECON) ____________________

Assoc. Prof. Dr. Turan EROL (CMB) ____________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Altan ALDAN

Signature:

ABSTRACT

CONVERGENCE ACROSS PROVINCES OF TURKEY:
A SPATIAL ANALYSIS

Aldan, Altan
M.Sc., Department of Economics
Supervisor: Assist. Prof. Dr. Esma Gaygısız

July 2005, 69 pages

The aim of this study is to analyze regional disparities and to test the convergence hypothesis across the provinces of Turkey. The study also attempts to analyze the spatial spillovers in the growth process of the provinces. The analyses cover the 1987-2001 period. Two alternative methodologies are used in the analyses. First, the methodology of β-convergence based on cross-sectional regressions is used and effects of spatial dependence are analyzed using spatial econometric techniques. Second, Markov chain analysis is used and spatial dependence is integrated using spatial Markov chains. Results of both methodologies signal non-existence of convergence and existence of spatial spillovers in the growth process of provinces.

Key Words: Regional Disparities, β-convergence, Markov Chains, Spatial Econometrics.

ÖZ

TÜRKİYE'NİN İLLERİ ARASINDA YAKINSAMA: MEKANSAL BİR ANALİZ

Aldan, Altan
Yüksek Lisans, Ekonomi Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Esma Gaygısız
Temmuz 2005, 69 sayfa

Bu çalışmanın amacı Türkiye'de bölgesel eşitsizliklerin analiz edilmesi ve yakınsama hipotezinin Türkiye'deki iller için test edilmesidir. Çalışma ayrıca, illerin gelişme sürecinde mekansal yayılma etkilerini analiz etmektedir. Çalışma 1987-2001 yıllarını kapsamaktadır. Analizlerde iki farklı yöntem kullanılmıştır. İlk olarak, yatay kesit regresyonlarına dayanan β-yakınsama yöntemi kullanılmış ve mekansal bağımlılığın etkileri mekansal ekonometri araçları kullanılarak bulunmuştur. İkinci olarak, Markov zincirleri yöntemi kullanılmış ve mekansal bağımlılık mekansal Markov zincirleri kullanılarak analize dahil edilmiştir. Her iki yöntemin sonuçları Türkiye'de iller arasında yakınsama olmadığına ve mekansal yayılmanın büyüme sürecinde etkili olduğuna işaret etmektedir.

Anahtar Kelimeler: Bölgesel Eşitsizlikler, β-yakınsaması, Markov Zincirleri, Mekansal Ekonometri

ACKNOWLEDGEMENTS

I would like to thank Assoc. Prof. Dr. Esma GAYGISIZ for her valuable comments, guidance and help in the preparation of this thesis.

I would like to thank Assoc. Prof. Dr. Turan EROL of Capital Markets Board (CMB) and Prof. Dr. Erol TAYMAZ for their participation in my examining committee and for their precious comments and suggestions.

I am grateful to Onur ÇAMLAR for his technical assistance. I wish to express my appreciation for the support of my colleagues in the Research Department of the Central Bank of the Republic of Turkey, especially M. Eray YÜCEL, Tayyar BÜYÜKBAŞARAN and Çağrı SARIKAYA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Regional disparities have been one of the most fundamental problems in Turkey for years. Reducing gaps in income and standard of living between rich West and poor East has become an important issue in politics and economic policy making. Since 1970s, five-year development plans include a regional perspective. Some regional development programs like Southeastern Anatolia Project (GAP), Eastern Anatolia Project (DAP) and Eastern Black Sea Project (DOKAP) have been developed and implemented to improve the socio-economic conditions in the lagging provinces in these regions. Additionally, investment incentives have been used to promote private investment and economic development in the least developed provinces.

Lack of public infrastructure investments in the least developed regions, which are crucial to promote private sector manufacturing by positive externalities, affect regional development negatively in these regions. Another obstacle in these provinces is the lack of educated work force. Only 13 of 74 universities are located in the least developed Black Sea, Eastern Anatolia and South Eastern Anatolia regions. These universities have severe shortages of human resources and physical infrastructure (DPT 2000).

Regional disparities are one of the determinants of migration within the country. Migration from least developed provinces to metropolitan areas like İstanbul, İzmir, Ankara and Adana, cause severe social and economic problems such as the inadequacy of education, health, infrastructure, and high unemployment in these cities (DPT 2000).

Reducing income gaps has also been an important policy issue in the European Union (EU) as well as in Turkey. The objective of reducing disparities across regions in the EU is laid down in the preamble to the Treaty of Rome (1957). After inclusion of Greece, Spain and Portugal, this objective has been further emphasized and annual spending on regional policy has increased (Neven and Gouyette 1995). Regional Development Fund comprises almost half of the structural funds in EU (DPT 2000).

In line with the increasing importance in politics and economic policy making, whether countries and regions converge in terms of per capita income or output has become one of the prominent issues in the literature. Tests of convergence in income are also used to assess alternative growth theories. Neoclassical growth theory pioneered by Solow (1956) concludes that there will be convergence in per capita income in the long run across economies (the term economy is used in the literature to represent both countries and/or regions depending on the study), which have the same steady state income level. Proponents of neoclassical growth theory have tried to show the existence of convergence process whereas opponents have tried to refute their findings and show that there is no clue for convergence of economies to a common steady state per capita income.

The objective of this study is to investigate whether convergence process has occurred across provinces of Turkey in the period from 1987 to 2001. The study uses two different methodologies: traditional approach and distribution dynamics approach. The traditional approach examines whether initially poor regions grow faster than the initially richer ones. Distribution dynamics approach examines the changes in cross section distributions of per capita income over time.

The main focus of the study is to analyze the effects of spatial dependence between provinces of Turkey in the growth and the convergence process. Since, it is unrealistic to assume regions within a country as independent of each other, recent studies on convergence issues take spatial dependence into account.

Spillover effects between provinces are calculated and spatial dependence is integrated both in traditional approach and distribution dynamics approach.

The study is organized as follows. Chapter II reviews the empirical models that analyze convergence. Chapter III deals with testing spatial dependence and integrating it in the convergence analysis. Chapter IV applies the alternative methodologies to test convergence in Turkey and integrates spatial dependence in the analysis. Finally, Chapter V derives the main conclusions.

CHAPTER II

LITERATURE SURVEY ON CONVERGENCE EMPIRICS ACROSS ECONOMIES

Regional disparities and income convergence are extremely important in policy making. There are two main approaches to assess regional disparities, growth and convergence within countries. First approach argues that, developments in transport and communications help reduce regional disparities since lagging regions have cost advantage due to cheap labor. Therefore, there is no need for special policies to reduce regional disparities. On the other hand, second approach argues that the fastest growing activities such as high technology industries and business-services are mainly concentrated in the most developed regions. Furthermore, policies emphasizing competitiveness increase agglomeration due to positive externalities and thus increase regional disparities. Therefore, policies to reduce regional disparities must be implemented (Gezici and Hewings 2001).

The relationship between national growth and regional convergence is also important for policy makers. Williamson (1965) argues that the typical pattern of national development creates regional divergence in the early stages of development and regional convergence in later stages. The main argument for this result is that growth in developing countries is generated by a limited number of growth poles, which enjoy the positive effects of agglomeration.[1] Therefore, for developing countries, growth of national income will increase regional disparities and the two goals of economic policy, i.e., reducing gaps between regions and maximizing national growth, may be conflicting.

[1] See Davies and Hallet (2002) for details of Williamson hypothesis.

After the seminal works of Baumol (1986) and Barro and Sala-i Martin(1991), convergence in per capita income across countries and within countries have become one of the most prominent issues in empirical economics. Following these papers, a large number of studies tried to uncover whether there is convergence among or within countries.

The theoretical background for the first empirical studies of income convergence was the neoclassical growth theory formulated by Solow (1956), which implies that all economies will converge to balanced growth paths with constant capital per effective labor, regardless of their initial conditions. Solow model is investigated in section II.1.

Barro and Sala-i Martin show that, under certain conditions, the process of convergence will also apply in per capita incomes and economies with initially lower per capita incomes will grow faster. Therefore, if a significant negative relationship between initial per capita incomes and growth rates of economies are found, it is argued that convergence exist and neoclassical growth theory is valid to explain growth. The methodology of Barro and Sala-i Martin is examined in detail in section II.2.

There have been many criticisms to the methodology of Barro and Sala-i Martin. Some of the criticisms have claimed that empirical finding of convergence in this methodology does not validate neoclassical growth theory, some criticize the lack of spatial effects in the regressions and some of the critics reject the use of regressions wholly. Criticisms to the traditional methodology are examined in section II.3. Quah (1993b) proposes a distribution dynamics approach, which examines the evolution of per capita incomes in time. This approach is examined in details in section II.4. Finally, section II.5 will conclude.

II.1 Basic Theoretical Growth Model: Solow Model

The Solow model focuses on four variables: output (*Y*), capital (*K*), labor (*L*) and "knowledge" or the "effectiveness of labor" (*A*). It is assumed that, labor and knowledge grow exponentially at the exogenous growth rates of n and μ, respectively. That is

$$A(t) = A(0)e^{\mu t} \text{ and } L(t) = L(0)e^{nt} \qquad \text{II.1}$$

Output is divided between consumption and investment. An exogenous fraction, s, of output is devoted to investment, which is used both for adding new capital and for replacing the existing capital which depreciates at an exogenous rate of δ. Therefore, the change of capital in time can be written as

$$\dot{K}(t) = sY(t) - \delta K(t) \qquad \text{II.2}$$

where $\dot{K}(t)$ denotes the time derivative of capital stock.

The production function, suppressing subscripts t denoting time, can be written as

$$Y = F(K, AL) \qquad \text{II.3}$$

Technological change in this model is known as labor-augmenting since A and L enter the production function multiplicatively. The term *AL* in the production function is called effective labor.

The model assumes that the production function has constant to returns in its two arguments, jointly. The production function is twice differentiable in its arguments, increasing and strictly concave. Dividing both sides by effective labor, the production function can be written in intensive form as

$$y = f(k) \qquad \text{II.4}$$

where y and k are output and capital per effective labor; $y = Y / AL$, $k = K / AL$. Output per effective labor is an increasing and concave function of capita per effective labor. The above assumptions imply that the production function in intensive form has decreasing returns to scale.

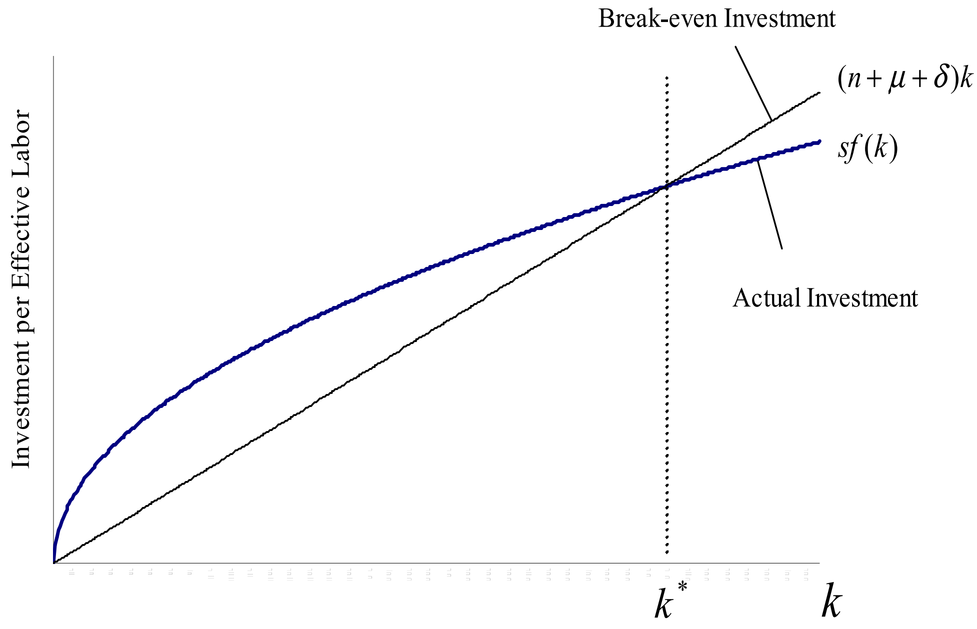
The time derivative of k can be written as (using chain rule),

$$\dot{k}(t) = \frac{\dot{K}(t)}{A(t)L(t)} - \frac{K(t)}{A(t)L(t)}\frac{\dot{L}(t)}{L(t)} - \frac{K(t)}{A(t)L(t)}\frac{\dot{A}(t)}{A(t)} \qquad \text{II.5}$$

Using (II.2), (II.5) and the fact that the growth rates of A and L, which are given exogenously as μ and n, respectively, we can show that

$$\dot{k}(t) = sf(k(t)) - (n + \mu + \delta)k(t) \qquad \text{II.6}$$

Figure 1 plots the two components of $\dot{k}$ as functions of *k*. For small values of *k*, the $sf(k)$ line, denoting the actual investment, is larger than the $(n + \mu + \delta)k$ line, representing the break-even investment. Therefore, for small values of *k*, actual investment is larger than break-even investment and capital stock per effective labor increases. As *k* gets larger, the slope of actual investment line decreases and falls below the slope of the break-even investment line and two lines eventually cross.



Source: Romer (2001).

Figure 1. Actual and break-even investment in Solow Model

As a summary, if k is initially less than k^*, actual investment is higher than the break-even investment and capital per effective labor is rising. On the other hand, if k is higher than k^*, then capital per effective labor is decreasing. If k equals k^* then there is no change in capital per effective labor. Therefore, regardless of where k starts, it converges to k^*. At the steady state where $k=k^*$, capital stock grows at the rate of $n+\mu$ since and labor and technology grow at the rates of n and μ, respectively. Since both inputs, capital and effective labor, grow at the rate of $n+\mu$, output also grows at $n+\mu$. Therefore, Solow model implies that the economy tend to converge to a steady state with a balanced growth path where each variable of the model grow at a constant rate.

The convergence scheme of Solow model has important implications about income differences across economies. First, the model predicts that all economies will converge to their balanced growth paths. To the extend that differences in output per worker arise from countries being at different initial points relative to their steady states, poorer countries are expected to grow faster than the richer ones. On the other hand, the assumption of decreasing returns to scale of capital per effective labor implies that capital is more effective in poor countries than the richer ones. Thus, there are incentives for capital to flow from rich to poor countries.

II.2 Traditional Approach to Income Convergence

Barro and Sala-i Martin (1991) develop the idea of income convergence using the implications the Solow model, which formed the traditional approach to income convergence. The traditional approach deals with the differences between growth rates of economies and concludes that there is income convergence if poorer economies grow faster thant the richer ones.

II.2.i Convergence Concepts and Methodology

Barro and Sala-i Martin (1991) use Cobb-Douglas production function with constant returns to scale and show that convergence process of capital per effective labor also applies for output per capita. The Cobb Douglas production function in intensive form can be written as

$$y = f(k) = k^{\alpha} \qquad \text{II.7}$$

where $0<\alpha<1$. In this production function, the growth rate of capital per effective labor can be written as

$$\frac{\dot{k}}{k} = sk^{-(1-\alpha)} - (n + \mu + \delta) \qquad \text{II.8}$$

A log-linear approximation of equation II.8 around the neighborhood of steady state yields

$$\frac{\dot{k}}{k} = \frac{d[\log(k)]}{dt} \cong -\beta[\log(k/k^*)] \qquad \text{II.9}$$

where β determines the speed of convergence from k to k^*. β is calculated as[2]

$$\beta = (1-\alpha)(n+\mu+\delta) \qquad \text{II.10}$$

In the Cobb-Douglas production function, we have

$$\frac{\dot{y}}{y} = \alpha\frac{\dot{k}}{k}$$

$$and \qquad \text{II.11}$$

$$\log(y/y^*) = \alpha\log(k/k^*)$$

Substituting II.11 to II.9 yields

$$\frac{\dot{y}}{y} = \frac{d[\log(y)]}{dt} \cong -\beta\left[\log\left({}^{y}\!/\!_{y^*}\right)\right] \qquad \text{II.12}$$

Equation II.12 is a differential equation with the solution

$$\log[y(t)] = \log[y(0)]e^{-\beta t} + \log(y^*)(1-e^{-\beta t}) \qquad \text{II.13}$$

---

[2] See appendix of chapter 1 in Barro and Sala-i Martin (1995) for derivation of equations II.9 and II.10

Then, the average growth rate of per capita income, $\hat{y}$ over the interval between dates 0 and T is

$$\frac{1}{T}\log\left[\frac{\hat{y}(T)}{\hat{y}(0)}\right] = \mu + \frac{1-e^{-\beta T}}{T}\log\left[\frac{y^*}{y(0)}\right]. \qquad \text{II.14}$$

Higher the convergence parameter β and higher the gap between the y^* and $y(0)$, then higher is the average growth rate of $\hat{y}$. Therefore, if two economies have the same μ and y^* then the economy with lower initial per capita income, $\hat{y}(0)$, will grow faster. Hence, if negative relationship between growth rates and initial per capita incomes are found, β-convergence is said to occur. In order to test convergence among economies *i* to *N*, growth rates of economies are regressed to their initial per capita incomes as

$$\frac{1}{T}\log\left(\frac{\hat{y}_{i,t_0+T}}{\hat{y}_{i,t_0}}\right) = B - \left(\frac{1-e^{-\beta T}}{T}\right)\log\left(\hat{y}_{i,t_0}\right) + u_i \qquad \text{II.15}$$

where *B* is a constant term and u_i are error terms which are assumed to satisfy standard Gauss-Markov assumptions. This type of convergence is called as unconditional or absolute convergence.

The assumption in absolute convergence framework that all economies have the same preferences, tastes and same steady states is quite restrictive and unrealistic. In the real world, economies may differ in their levels of technology, propensities to save, rates of population growth etc. Sala-i Martin (1996) argues, in his own words, that

> *Because we think that the technology, institutions and tastes of most African economies are very different from those of Japan or United States,*

the assumption that these economies converge to a common steady state is not realistic

Two different ways to maintain the steady state constant are used. The first one is to restrict the data set with economies that are thought to have similar steady states. The technological and institutional differences across regions within a country or across 'similar' countries are probably smaller (Sala-i Martin 1996). Therefore, if neoclassical growth theory is valid then there must be absolute β-convergence across regions in a country.

The second way to hold steady state constant is to introduce control variables that proxy the steady states of different economies to the β-convergence equation in II.15. This type of convergence is called conditional β-convergence. If the coefficient of initial income level is negative once the steady state is held constant via the control variables then conditional β-convergence occurs.

The effect of the initial per capita income on the average rate of growth gets smaller as T gets larger. Therefore, β is estimated non-linearly to take account T so that similar estimates of β are obtained regardless of the time interval. If estimated β in this specification is significantly positive then existence of convergence (initially poorer economies growing faster than the richer ones) is concluded.

Another convergence concept commonly used in the traditional literature is σ-convergence developed by Baumol (1986). Although it has nothing to do with the neoclassical growth model, it has generally been used by researchers in traditional approach as a complement to β-convergence. There is σ-convergence if the dispersion of per capita income across the weighted-mean declines over time. In general, standard deviation or coefficient of variation is used as a measure of dispersion.

Concepts of β-convergence and σ-convergence are not identical, though related. The former relates to the mobility of per capita income within the same distribution whereas the latter relates to the evolution over time of the distribution of per capita income. Unconditional β-convergence is a necessary but not a sufficient condition for σ-convergence (Barro and Sala-i Martin, 1991).

II.2.ii Empirical Findings

The study of Baumol was the first pioneering study that gave rise empirical studies of convergence. In his study of productivity growth, Baumol concluded that the finding of negative correlation between initial productivity level and average productivity growth in a cross section of the states of US in the period 1870-1979 implies that there is convergence in productivity levels between states of US in the sense that states initially with low productivity levels catch up with the states with initially high levels of productivity.

Barro and Sala-i Martin (1991) examine convergence across the US states. Using data for per capita personal income, exclusive of all transfers, for 47 states for the period between 1880-1988 they find evidence of convergence in the US. They divide the period into 9 sub periods and find evidence of convergence except for two sub periods. For the whole period they get the convergence speed about 2% per year. Using the data for gross state product for 48 states for the period 1963-1986 they obtain similar results. They also find evidence of convergence examining data for the regions of Germany, United Kingdom, Italy, France, Netherlands, Belgium and Denmark for the period 1950-1985. They conclude that there is convergence both within countries and between countries.

Barro (1991) uses the Summers-Heston (1988) data set to analyze convergence of 98 countries from 1960 to 1985. He finds, in absolute terms without additional variables to hold steady state constant, a negative convergence coefficient meaning that rich countries grow faster and that data exhibits divergence. Since steady states of the countries in the sample are far from being similar, he adds control variables. The set of control variables consists of primary and secondary school enrollment rates in 1960, the average ratio of government consumption expenditure (exclusive of defense and education) to GDP from 1970 to 1985, proxies for political stability and a measure of market distortions based on purchasing power parity ratios for investment goods. After inclusion of these control variables, he finds conditional convergence with a speed of 2 %. He obtains similar results for 20 OECD countries, as well.

Sala-i Martin (1996) finds evidence of absolute convergence in Japanese prefectures in the period 1955-1987. The estimated rate of convergence is again 2 %. He also reports absolute convergence within five European countries, Italy, UK, Germany, France and Spain with speed of convergence ranging from 1.5 % to 2.9 % in the period from 1950 to 1990. For the same time period, he also finds conditional convergence in European regions using country dummies with a convergence rate of 1.5 %. Therefore, he concludes that as a general rule, there is convergence with a speed of around 2 %.

Neven and Gouyette (1995) find evidence of conditional β convergence in 141 European regions at NUTS (Nomenclature of Territorial Units for Statistics)[3] II level in the period 1980-1989. However, the rate of convergence is very low compared to the finding of Sala-i Martin. De La Fuente (2002) reports absolute β convergence in Spain in the period of 1965-1995. The rate of convergence, however, declines from 2.49 % in the period 1965-1975 to 0.38 % in the period 1985-1995. Michelis *et al.* (2004) examine the convergence process in Greece in the period of 1981 and 1991, just after the entrance of Greece to European Union and find that there is absolute convergence in Greece in that period with a

[3] Nomenclature of Territorial Units for Statistics (NUTS) is the standard for referencing administrative division of countries for statistical purposes in European Union. There are three levels of NUTS. NUTS 1 describes the broader regions and NUTS 3 describes smaller ones. In Turkey, State Institute of Statistics collects some data on the basis of NUTS classification. Provinces correspond to NUTS3 level of classification.

convergence parameter of 1 %. Kosfeld *etal.* (2002) examine the convergence process in Germany after unification. They estimate an absolute convergence parameter of 6.5 % per year in the period of 1992-2000.

Several studies have examined the convergence process in Turkey using traditional approach. Tansel and Güngör (1998) construct a provincial labor productivity series by dividing provincial GDP (at 1987 prices) by labor forces of each province for the period 1975-1995. Using β-convergence regressions, they obtain absolute convergence in labor productivity across provinces of Turkey at a rate of 0,2 % per year for the period 1975-1995 and 0,5 % for the period 1980-1995. Erk et al (2000) conclude that there is no evidence of convergence in the provinces of Turkey in GDP per capita for the period 1979-1997 using σ and β-convergence analysis. Gezici and Hewings (2002) and Karaca (2004) obtain similar results.

II.3 Criticism on Traditional Approach to Convergence

The concepts of both unconditional and conditional β-convergence and the general empirical finding that economies converge to their steady states at a rate of 2 % have been forcefully criticized in the recent literature. In case of unconditional β-convergence, all economies converge to the same steady-state. In case of conditional β-convergence, an economy approaches to its own but unique, globally stable, steady-state equilibrium. Chatterji (1992) proposed the notion of club convergence, which does not necessitate the existence of a unique steady state. A convergence club is a set of countries or regions for which growth rates and initial per capita incomes are negatively correlated. The concept of club convergence is based on endogenous growth models that are characterized by the possibility of multiple, locally stable, steady state equilibria as in Azariadis and Drazen (1990). Which of these different equilibria an economy will be reaching depends on the range to which its initial conditions belong. If convergence clubs exist, β-convergence equations can lead to the conclusion of validity of neoclassical growth model although it is not actually the case. Therefore, β-convergence equations cannot distinguish between neoclassical growth models and endogenous growth models.

Durlauf and Johnson (1995) use Summers-Heston data set and detect the existence of convergence clubs in the sample of 121 countries for the period 1960-1985. They split the data in terms of control variables and check whether sub samples behave differently. They estimate different coefficients in the convergence regressions when the data is divided by initial income and literacy rates. Therefore, they conclude the presence of convergence clubs. Chatterji and Dewhurst (1996) conclude that there exist two convergence clubs in Great Britain using a nonlinear specification by adding higher powers of initial per capita income as additional regressors.

Another issue of criticism of traditional β-convergence studies is the negligence of the spatial spillovers. The assumption of spatial independence of economic activities between countries may somehow be defended. However, it is totally unrealistic to neglect spatial dependence in the regional studies within a country where factors of production are more mobile. Developments in spatial econometrics propagated the usage of models taking into consideration the spatial dependence. However, in the traditional β-convergence studies the spatial econometric techniques were not used. The spatial effects were only handled by simply using regional dummies as in Barro and Sala-i Martin (1991) for the US states. Although this specification can be useful to reduce or totally eliminate the spatial autocorrelation in the error terms, it is restrictive and does not give information about spatial spillovers within or across countries.

Some recent studies use spatial econometric techniques in convergence equations. Rey and Montuori (1999) find strong evidence of positive spatial dependence among 48 US states in both levels and growth rates of per capita income for the period 1929-1994. They conclude that traditional unconditional β-convergence model suffers from spatial dependence. Baumont *etal*. (2002), using GDP per capita data for 138 European regions at NUTS I level, conclude that there is spatial autocorrelation in β-convergence regression. They also estimate a strong spatial spillover effect and propose that average growth rate of per capita GDP of a given region is positively affected by the average growth rate of neighboring regions.

The most radical criticisms were however directed to the use of regression-based techniques to test the convergence hypothesis. It is pointed out that regressions concentrate on the behavior of the representative economy that can give information on the transition of this economy towards its own steady state whilst giving no information on the dynamics of the entire cross-sectional distribution of income. Friedman (1992) and Quah (1993a) demonstrate that negative coefficient of initial income can be associated with divergence as well as convergence. Therefore, coefficient of initial income says nothing about whether there is convergence or divergence. The fact that negative coefficient of initial income is insufficient for convergence is also acknowledged by the proponents of β-convergence equations. They argue that a negative coefficient must be interpreted as indicating the existence of forces reducing the cross-sectional distribution while ongoing shocks have reverse effects. Therefore, β-convergence must be jointly interpreted with σ-convergence.

Quah (1993b) agrees with the idea that σ-convergence is more informative than the β-convergence equations. However, he argues that using σ-convergence reduces the evolution of cross-sectional distribution of income to a single statistic. Instead, convergence should be studied by taking into account the shape of the entire distribution of per capita GDP and its intra-distribution dynamics over time and not by estimating the cross section correlation between growth rates and per capita GDP levels. His study leads to the alternative methodology of convergence based on distribution dynamics.

II.4 Distribution Dynamics Approach to Convergence

After the seminal work of Quah, many studies used his methodology to analyze convergence, which led to distribution dynamics approach. Distribution dynamics approach deals with the cross sectional distributions of per capita incomes and the evolution of these distributions over time.

II.4.i Methodology

Let F_t denote the cross-sectional distribution of per capita incomes at time t. Then the evolution of this distribution over time can be described by the following equation

$$F_{t+1} = T(F_t) \qquad \text{II.16}$$

where *T* is an operator that describes the transition from one distribution into the other.

Two ways of analyzing convergence in the framework of equation II.16 is possible. The first one is to treat F_t as continuous. Then, a probability distribution is estimated for F_t and the operator becomes *T* can be interpreted as a stochastic kernel (Quah 1996a). Convergence is analyzed using the shapes of F_t and analyzing the shape of three-dimensional plot of the stochastic kernel.

The second way to analyze convergence is to treat income space as discrete. Then, F_t can be represented by probability vectors and the operator T becomes a probability transition matrix, P. In that case, equation II.17 can be rewritten as

$$F_{t+1} = P.F_t \qquad \text{II.17}$$

and the system is treated as a first-order Markov process.

Using stochastic kernels has advantage over using discrete Markov chains in the sense that there is some arbitrariness in discretization. On the other hand, while stochastic kernels allow characterizing the evolution of global distribution they do not provide any information about the movements of the regions within this distribution (Le Gallo 2004). Therefore, while stochastic kernels are not as restrictive as discrete Markov chains, they are not as informative as discrete Markov chains as well. In this study, discrete Markov chains will be used to analyze convergence.

The analysis of Markov chains starts with defining a set C of K income classes (or states). F_t becomes the probability vector of these classes at time t, that is $F_t = (F_{1t}, F_{2t}, ...F_{Kt})'$. Then P can be interpreted as a transition probability matrix: for any two income classes i and j ($i, j \in$ C), the element p_{ij} of P define the probability of moving from class i to class j between time t and $t+1$. (Magrini 2004). In that case, a (first-order, discrete) Markov chain is defined as a stochastic process such that, for any variable x of a region r, the probability p_{ij} of being in a state j at any point of time $t+1$ depends only on the state i it has been at t, but not on the states at previous points of time, that is (Bickenbach and Bode 2003)

$$P\{x_{r,t+1} = j | x_{r,t} = i, x_{r,t-1} = i_{-1}, ..., x_{r,0} = i_0\} = P\{x_{r,t+1} = j | x_{r,t} = i\} = p_{ij} \quad \text{II.18}$$

for any region r and for any $i, j \in$ C. Equation (2) is usually referred to as Markov property.

If the process is time independent, the Markov chain is completely determined by the Markov transition matrix P with $p_{ij} \geq 0$ and $\sum_j p_{ij} = 1$ which summarizes all

K^2 transition probabilities and an initial distribution $h_0 = (h_{1,0}, h_{2,0}, \ldots\ldots h_{K,0})$, $\sum_i h_{i,0} = 1$ describing the starting probabilities of the various states.

Important information about the dynamics of the cross-sectional distribution can be obtained by considering higher order transition probabilities $p_{ij}(l)$. In this case, the transition matrix $P(l)$ contains information about the probabilities that take place in exactly l periods. Higher order transition probabilities have the relationship (Chapman-Kolmogorov Equation)

$$p_{ij}(l_1 + l_2) = \sum_m p_{im}(l_1) p_{km}(l_2) \qquad \forall i, j, m \in C \qquad \text{II.19}$$

In terms of the transition matrices, Chapman-Kolmogorov equation can be written as

$$P(l_1 + l_2) = P(l_1) P(l_2) \qquad \text{II.20}$$

It is also informative to find the limiting probabilities of states in the long run, p_i $i \in$ C. However not all Markov chains have limiting probabilities. If a Markov chain is ergodic it has a limiting (stationary, ergodic) distribution.[4]

The transition matrix can be estimated by a Maximum Likelihood approach (Bickenbach-Bode 2001). Assume that there is only one transition period, with the initial distribution $h_i = n_i / n$ being given and let n_{ij} denote the empirically observed absolute number of transitions from i to j. Then, maximizing

$$\ln L = \sum_{i,j=1}^{K} n_{ij} \ln p_{ij} \qquad \text{s.t.} \sum_{\mathrm{j}} p_{ij} = 1,\ p_{ij} \geq 0 \qquad \text{II.21}$$

[4] Conditions for ergodic Markov chains are discussed in detail in the Appendix.

with respect to p_{ij} gives

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{\mathrm{j}} n_{ij}} \qquad \text{II.22}$$

as the asymptotically unbiased and normally distributed Maximum Likelihood estimator of p_{ij}.

The reliability of the Markov transition probabilities depends on the assumption of homogeneity over time, which means that the transition probabilities do not change over time. In order to test time homogeneity, whole period is divided into sub periods and the hypothesis that transition probabilities estimated for sub periods do not differ than those estimated for the entire period. In order to test the hypothesis, the following test statistic is utilized (Bickenbach Bode 2003)

$$Q^{(T)} = \sum_{t=1}^{T}\sum_{i=1}^{K}\sum_{j=1}^{K} n_i(t)\frac{(\hat{p}_{ij}(t) - \hat{p}_{ij})^2}{\hat{p}_{ij}} \qquad \text{II.23}$$

where $\hat{p}_{ij}$ is the probability of transition from class i to class j estimated from the whole period, $\hat{p}_{ij}(t)$ is the corresponding transition probability estimated from sub period t and $n_i(t)$ is the number of observations in class i in sub period t. T is the number of sub periods and K is the number of classes. The statistic is distributed as χ^2 with degrees of freedom of $\sum_{i=1}^{K}(a_i - 1)(b_i - 1)$ where a_i is the number of sub periods in which observations for the *i-th* row are available and b_i is the number of positive entries in the *i-th* row of the matrix for the entire sample.

Analysis of convergence is done by examining the probabilities p_{ij} and the ergodic distribution. If the probability of moving to richer classes is high in the poor income classes, then convergence is said to occur since a region starting from a poor income class have the chance to become richer. If the probability of middle income classes is higher than the probabilities of classes in the tails of the distribution in the ergodic distribution compared with the initial distribution, again convergence is concluded. On the other hand, if the ergodic distribution is concentrated around two distinct classes, then formation of convergence clubs or bimodality in the income distribution is concluded.

II.4.ii Empirical Findings

Quah (1993b) is the first study that analyses convergence in the distribution dynamics framework. He uses empirical methodology based on Markov chains by partitioning the income space and tries to examine the change in income distribution over time. For the GDP per capita data for 118 countries in the period 1962-1985, he finds persistence of economies in their initial states and concludes that there seems to be no sign of convergence.

Quah (1996b) finds evidence of convergence of US states since the transition probabilities in the Markov chain reveal a high degree of mobility among classes and the ergodic distribution presents no sign of bimodality. Johnson (2000) supports the findings of Quah(1996b) using stochastic kernel estimation procedure for transitions of distributions.

Neven and Gouyette (1995) employ Markov chains to analyze convergence and find no evidence of convergence in the European Community using GDP per capita data for 141 European regions at NUTSII level in the period 1980-1989. Low mobility observed in the poorest income classes suggests that poorest regions in the Community are very likely to stay poor. Lopez-Bazo *et al.* (1999) support the findings of Neven and Gouyette in their study using GDP per worker data for 129 European regions in the period 1983-1992. Magrini (2004) employs stochastic kernels to examine convergence in 110 NUTSII regions in the European Union in the period of 1980-1995. He also concludes that there exists no evidence of convergence in the regions of the European Union. Pekkala (2000) finds high mobility and a potential of convergence in his study using GDP per capita data for 88 Finnish sub-regions in the period 1988-1995.

To the best of my knowledge, the only study that deals with convergence issue using Markov chains is that of Temel *et al.* (1999). They use labor productivity data for provinces of Turkey in the period 1975-1990. They conclude, in terms of labor productivity, in contrast to the study of Tansel and Güngör (1998), which used traditional approach, that there seems to be no evidence of convergence. Indeed, middle classes tend to diminish and convergence clubs tend to emerge in the time invariant distribution.

II.5 Conclusion

Convergence in per capita income has been subject of many empirical studies in the last decade. Besides being an important policy issue, convergence attracted this much interest since it was seen as a tool to test growth models. First studies of convergence implicitly attempted to test the validity of neoclassical growth model (the Solow model). Early findings of existence of convergence in states of US and across countries were criticized and different methodologies are utilized to test convergence.

The empirical results differ according to the choice of space and period to be studied. In almost all studies related with US, existence of convergence is concluded. On the other hand, different results are obtained in the studies for European Union. Different periods, different samples and different methodologies result with conflicting results. The studies about Turkey find no significant evidence of convergence in Turkey.

As pointed out in section II.3, spatial effects are neglected in the early studies of convergence. However, spatial dependence is expected especially in the regional data within a country. Studies that do not consider spatial effects may give misleading results. And space may explain the growth of regions to some extend. Therefore, space must be integrated into the convergence analyses. In the next chapter, spatial dependence will be examined and will be integrated to the convergence analyses.

CHAPTER III

SPATIAL DEPENDENCE IN CONVERGENCE EMPIRICS

Regional studies use spatial data, which have special properties and need to be analyzed in different ways from nonspatial data. Spatial effects, most importantly spatial dependence must be included in the regional analyses although studies of convergence have generally neglected these effects. Several factors, such as trade between regions, technology and generally spatial spillovers may cause to geographically dependent regions. Negligence of spatial autocorrelation in regional data may cause misleading results. Therefore, when dealing with regional data, existence of spatial autocorrelation must be explored. If there is spatial autocorrelation in the data under study, then an appropriate model that will take it into account must be used.

Section III.1 will give the formal definition of spatial autocorrelation and define how to measure it. Section III.2 will integrate spatial dependence in convergence studies first in traditional approach and then distribution dynamics approach. Section III.3 will give some examples of studies on convergence that take spatial dependence into account. Section III.4 concludes.

III.1 Measuring Spatial Dependence

Spatial dependence in a sample refers to the fact that one observation associated with a location i depends on other observations at locations j $(j \neq i)$ That is,

$$x_i = f(x_j) \qquad i = 1,.........N \qquad j \neq i \qquad \text{III.1}$$

where x is the variable under consideration. Two broad sources of spatial dependence are generally pointed out. First, it is a result of spatial interaction effects such as technological spillovers and factor mobility. Second, it may be due to the measurement problems resulting from the fact that administrative borders may not coincide with the borders of economic activity (Anselin 1988).

The most common statistic used for detecting the spatial dependence is the Moran's I statistic which is formulated as (Upton and Fingleton 1985)

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \qquad \text{III.2}$$

where n is the number of regions, S_0 is the sum of the elements in the spatial weight matrix W which summarizes the spatial effects between regions, w_{ij} are the elements of the spatial weight matrix W corresponding to the regions i and j. Moran's I statistic can take values between –1 and 1. Positive values of Moran's I indicate positive spatial autocorrelation in which similar values are more likely than dissimilar values between neighbors and vice versa. If x_i are distributed normally, then I can be assumed as normally distributed with expected value, $E(I)$ and variance, var(I) given as,

$$E(I) = -\frac{1}{n-1}$$

$$\mathrm{var}(I) = \frac{n\{(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2\} - k\{n(n-1)S_1 - 2nS_2 + 6S_0^2\}}{(n-1)^{(3)} S_0^2} - \frac{1}{(n-1)^2}$$

where $S_1 = \frac{1}{2}\sum_i \sum_j (W_{ij} + W_{ji})^2 \quad , (\mathrm{i} \neq \mathrm{j})$

$$S_2 = \sum_i (W_{i0} + W_{0i})^2 \text{ with } W_{i0} = \sum_j W_{ij} \text{ and } W_{0i} = \sum_j W_{ji} \,,$$

$$k = \frac{m_4}{m_2^{\,2}} \text{ with } m_r = \frac{1}{n}\sum_i (x_i - \bar{x})^r$$

$$\text{and } (n-1)^{(3)} = (n-1)(n-2)(n-3)\,. \qquad \text{III.3}$$

Spatial weight matrix is the fundamental tool to model and detect spatial dependence. Several forms of spatial weight matrices are suggested in the literature. Most commonly used weight matrix is the contiguity matrix having value of 1 if two regions i and j are neighbors and 0 for other entries of the matrix. Other forms of weight matrices are geographically based ones, i.e. those based on the inverse of distance between two regions, inverse of square of distance between two regions. Weight matrices based on population dynamics, agglomeration for example, and economic activities are also used. However, to avoid identification problems, the weight matrix based on purely spatial pattern must be used (Baumont *etal.* 2001). In this study, contiguity weight matrix is employed. To simplify calculations, the weight matrix is row standardized, that is sum of elements in each row is constrained to unity.

III.2 Integrating Spatial Dependence in Convergence Analysis

If there is spatial dependence in per capita incomes, then the convergence analysis may suffer from spatial autocorrelation. Earlier studies showed that, taking spatial dependence into account might change the results of the convergence analysis. Therefore, spatial dependence should be integrated to the convergence analysis. In this section, it will be integrated into convergence analyses first in traditional approach and then distribution dynamics approach.

III.2.i Traditional Approach

The empirical methodology of traditional approach to test convergence is based on cross-section regressions. In order to have correct results in these regressions, residuals must satisfy the standard Gauss-Markov assumptions. One of these assumptions is the independence of error terms. However, if there is spatial autocorrelation in the regional data, then the residuals of the regression may be spatially autocorrelated, which violates the Gauss Markov assumptions. In that case, the estimate of convergence parameter β will not be reliable. Therefore, existence of spatial autocorrelation in the residuals of the regression must be tested. A number of test statistics are suggested in the literature to test spatial autocorrelation in the residuals.

The most commonly used test statistic for spatial dependence in the residuals of a regression is the Moran's I statistic. It can be written as

$$I = \left(\frac{n}{S_0}\right)\frac{e'We}{e'e} \qquad \text{III.4}$$

where e is the vector of residuals from OLS regression, W is the spatial weight matrix, n is the number of regions and S_0 is the sum of all elements of the spatial weight matrix.

If the spatial weight matrix is row standardized, then the statistic simply takes the form

$$I = \frac{e'We}{e'e} \qquad \text{III.5}$$

The asymptotic distribution for Moran's I corresponds to a standard normal distribution after adjusting the I-statistic by subtracting the mean and dividing by the standard deviation (Anselin 1988). The mean and variance of the statistic can be written as

$$E(I) = \frac{\left(n/_S\right) tr(MW)}{n-k}$$

$$\mathrm{var}(I) = \left(n/_{S_0}\right)^2 \left[\frac{tr(MWMW') + tr(MW)^2 + (tr(MW))^2}{d}\right] - [E(I)]^2$$

with

$$d = (n-k)(n-k+2)$$

$$M = (I - X(X'X)^{-1}X') \qquad \text{III.6}$$

where X is the matrix of explanatory variables, tr denotes the trace operator, k is the number of explanatory variables.

A second test statistic is the likelihood ratio test that depends on the difference between the log likelihood from the spatial errors model, which will be discussed below and OLS regression. The statistic is distributed as χ^2 distribution with 1 degree of freedom.

Another approach is based on a Wald test for spatial dependence. The statistic can be written as (Anselin 1988)

$$Wald = \lambda^2 \left[t_2 + t_3 - \left(1/_n\right)(t_1{}^2)\right]$$

with

$$t_1 = tr(W.*B^{-1})$$

$$t_2 = tr(WB^{-1})$$

$$t_3 = tr\left[(WB^{-1})'(WB^{-1})\right] \qquad \text{III.7}$$

$$B = (I_n - \lambda W)$$

where $.*$ denotes element-by –element matrix multiplication and λ is the spatial correlation coefficient estimated in the spatial errors model. The statistic is distributed as χ^2 with 1 degree of freedom.

Final statistic based on least squares residuals to test the spatial dependence on OLS regression is the Lagrange Multiplier Test (LMERR). LMERR takes the form (Anselin 1988)

$$LMERR = \left(1/T\right)\left[(e'We)/\sigma^2\right]^2$$

$$T = tr(W + W').*W \qquad \text{III.8}$$

and is distributed as χ^2 distribution with 1 degree of freedom.

In all of the four tests discussed above the null hypothesis is non-existence of autocorrelation in the least squares residuals and large values of statistics lead to rejection of null hypothesis.

Several specifications are suggested in the existence of spatial autocorrelation in the error terms of an OLS regression. The easiest model used in the presence of spatial autocorrelation is the spatial cross-regressive model, which can be written as

$$Y = X\beta + WX\theta + u \qquad \text{III.9}$$

where Y contains an $nx1$ vector of dependent variables, X represents the nxk matrix of independent variables and W is the nxn spatial weight matrix summarizing the spatial effects between regions, WX is the spatial lag of the independent variable and u is the disturbance term satisfying usual Gauss-Markov properties. Since the spatial lag of the independent variable is exogenous, the model can be estimated via OLS.

In order to test whether spatial cross-regressive model eliminates spatial autocorrelation in the residuals, test statistics based on OLS residuals described above can be used since the model is estimated via OLS.

Another model is the spatial lag model (or spatial autoregressive model) where the spatial dependence is filtered out by the inclusion of spatial lag of dependent variable. The spatial lag model can be defined as (in vector form)

$$Y = X\beta + WY\rho + u \qquad \text{III.10}$$

where WY denotes the spatial lag of the dependent variable and the error terms u satisfy the Gauss Markov assumptions. Estimation of spatial lag model via OLS gives biased and inconsistent estimates. Consequently, maximum likelihood method is used to estimate the spatial lag model (Anselin 1988).

In order to examine whether the spatial lag model eliminates spatial autocorrelation, a Lagrange multiplier test based on spatial lag model (LMLAG) is used.

The test statistic is based on the model

$$Y = X\beta + CY\rho + u$$

$$u = \lambda Wu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where C is the spatial weight matrix of lagged dependent variable which may or may not be equal to W. In this study, weight matrix based on contiguity is used for C as well as W.

The test statistic is given by

$$LMLAG = \left(e'We \big/ \sigma^2 \right) \left[T_{22} - (T_{21})^2 \operatorname{var}(\rho) \right]^{-1}$$

$$T_{22} = tr(W.*W + W'W)$$

$$T_{21} = tr(W.*CA^{-1} + W'CA^{-1}) \qquad \text{III.11}$$

$$A = (I_n - \rho C)$$

where e represents the vector of residuals in of the spatial lag model and var(ρ) is the maximum likelihood estimate of the variance of the parameter ρ in the model.

The distribution of the statistic is χ^2 with 1 degree of freedom. The null hypothesis is absence of autocorrelation (λ=0) and high values of the statistic leads to rejection of null hypothesis of no autocorrelation in spatial lag model.

Spatial cross-regressive and spatial lag models are suitable to filter out spatial dependence that comes from spatial spillovers. On the other hand, if spatial dependence comes also from measurement problems, i.e. mismatch between borders of economic activity and administrative units, these models may be inappropriate. In such a case, the error term in the cross-section regression becomes non-spherical and spatial errors model is used, which can be defined as

$$\begin{aligned} &Y = X\beta + u \\ &u = \lambda W u + \varepsilon \qquad \text{III.12} \\ &\varepsilon \sim N(0, \sigma_\varepsilon^2 I_n) \end{aligned}$$

OLS estimate of β is unbiased but is inconsistent. Therefore, as in spatial lag model this model is also estimated via maximum likelihood method (Anselin 1988).

III.2.ii Distribution Dynamics Approach

Effects of spatial dependence have recently been included in the Markov chain analysis for convergence. The most informative method that shows the effects of spatial dependence in movements of regions within different income classes is the spatial Markov chain analysis suggested by Rey (2001).

In spatial Markov chain analysis, traditional Markov chain is modified in such a way that the transition probabilities of a province are conditioned on the class of its spatial lag for the beginning of the year. This procedure results in a transition matrix, which is a traditional *KxK* matrix decomposed into *K* conditional matrices of dimension *KxK*. Then an element in the *k-th* conditional matrix $\hat{p}_{ij}(k)$ gives the probability that a region in class *i* at time *t* moves to class *j* at *t+1*, given that its spatial lag is in class *k* at time *t*.

The spatial Markov chain allows one to examine the positive or negative influence of neighbors on the transition probability of a province. The influence of spatial dependence is reflected in the differences between the transition probabilities in the traditional Markov chain and conditional Markov chains. For example, in a spatial Markov chain with 5 classes (the poorest income class is class 1 and richest income class is class 5), for an upward transition from class i to class j ($i<j$), if the probability of transition in the unconditional traditional matrix p_{ij} is higher than the probability of transition conditional on spatial lag of 1 $p_{ij}(1)$, then provinces surrounded by poorest regions have lower opportunity to move up to an higher class. Conversely, if $p_{ij}(5)$, transition probability of a province surrounded by rich neighbors, is higher than p_{ij}, then provinces with richer neighbors are more likely to promote to a higher class.

To test existence of spatial dependence formally, a test statistic is developed by Bickenbach and Bode (2003). The null hypothesis of the statistic is the independence of the transition probabilities of space. In the null hypothesis, transition probabilities of a class in different spatial lags do not differ than each other and than the entire sample, that is

$$p_{ij}(1)= p_{ij}(2)= \dots p_{ij}(K)= p_{ij}$$

where K is the number of spatial lags which equals to the number of income classes. The test statistic can be written as

$$Q^{(R)} = \sum_{s=1}^{K}\sum_{i=1}^{K}\sum_{j=1}^{K} n_i(s)\frac{\left(\hat{p}_{ij}(s)-\hat{p}_{ij}\right)^2}{\hat{p}_{ij}} \qquad \text{III.13}$$

where $\hat{p}_{ij}$ is the probability of transition from class i to class j estimated from the whole sample, $\hat{p}_{ij}(s)$ is the corresponding transition probability estimated from sub sample with spatial lag s and $n_i(s)$ is the number of observations in class i in sub sample with spatial lag s. The statistic is distributed as χ^2 distribution with

$\sum_{i=1}^{N}(c_i - 1)(b_i - 1)$ degrees of freedom where c_i is the number of spatial lags in which the number of observations in class i is positive and b_i is the number of positive entries in the *i-th* row of the matrix for the entire sample.

III.3 Empirical Findings

Some recent studies take spatial effects into consideration. Rey and Montuori (1999) find strong evidence of positive spatial dependence among 48 US states in both levels and growth rates of per capita income for the period 1929-1994 and that traditional unconditional β-convergence model suffers from spatial dependence. They conclude that spatial errors model is the most appropriate model to examine convergence process in US.

Lopez-Bazo *etal* (1999) examine spatial dependence in 110 NUTS II level European regions for the period from 1983 to 1992. They find evidence of strong positive spatial autocorrelation for almost all years. Baumont *etal*. (2002), using GDP per capita data for 138 European regions at NUTS II level, conclude that there is spatial autocorrelation in β-convergence regression. They also estimate a strong spatial spillover effect and propose that average growth rate of per capita GDP of a given region is positively affected by the average growth rate of neighboring regions. Kosfeld *etal.* (2002) find strong spatial autocorrelation in per capita incomes of German regions. They also apply spatial econometric techniques in convergence analysis and conclude that the value of convergence parameter changes after including spatial dependence.

Le Gallo (2004) applies the method of spatial Markov chains to the sample of 138 European regions at NUTS II level. She finds evidence of strong spatial dependence. She proposes that, as the neighbors of a region get richer on average, the probability of the region to promote to a higher income class rises.

Gezici and Hewings (2002) use the Moran's I statistic to detect spatial dependence within Turkey and within regions of Turkey for the period 1980-1997. Regarding Turkey, they propose that the assumption of spatial dependence cannot be rejected in GDP per capita in the initial and end years, 1980 and 1997. On the other hand, they find no evidence of significant spatial autocorrelation in growth rate of per capita GDP between 1980 and 1997. Therefore, while the level of growth among provinces is dependent on the level of growth of neighbors, the growth rate seems to be more independent of the growth rate of neighbors.

III.4 Conclusion

Studies dealing with regional data have to consider spatial autocorrelation to obtain reliable results. On the other hand, literature on convergence generally did not consider the spatial aspects of regional data. Convergence studies that allow for the role of space are exceptions rather than norm. However, these studies show that spatial dependence has an important role in growth performances of the regions. In the traditional approach, the convergence parameters change significantly when spatial dependence is taken into account. In the distribution dynamics approach, transition probabilities conditional on spatial lags differ significantly from unconditional transition probabilities.

Gezici and Hewing (2002)'s finding of spatial dependence in per capita incomes of provinces in Turkey implies that spatial autocorrelation must be considered in studying convergence in Turkey.

In the next chapter, convergence will be analyzed. In doing so, first of all spatial dependence in levels and growth rates of per capita incomes of provinces will be analyzed. Then, alternative methods to test convergence discussed in chapter II will be used and spatial dependence will be integrated in these analyses.

CHAPTER IV

CONVERGENCE AMONG PROVINCES OF TURKEY AND EFFECTS OF SPATIAL DEPENDENCE

One of the main policy issues in Turkey has been reducing regional disparities especially between the provinces in the rich West and poor East. On the other hand, after 1980, export oriented policies emphasized the competitiveness in the international markets and gave privileges to the developed cities in the west of the country which are more competitive in the international markets. Therefore, two policies of decreasing regional disparities and maximizing international competitiveness have conflicting outcomes. Analyzing convergence will give some idea about whether the regional policies have been successful. This chapter tests convergence across provinces of Turkey using two alternative methods discussed in Chapter II.

In this study, GDP per capita is used as the measure of income to investigate convergence and spatial spillovers in the period from 1987 to 2001. Data for provincial GDP are taken from State Institute of Statistics (SIS) in 1987 constant prices. Population data are taken from official census done by SIS for the years 1985, 1990, 1997 and 2000. Population data for the years between the census years are interpolated. From 1989 to 1999 number of provinces in Turkey increased from 67 to 81. In this study, values related to 14 provinces established after 1989 were added to the values of the provinces from which they were separated for the sake of simplicity.

Empirical findings of Gezici and Hewings (2002) suggest that there is spatial dependence in per capita incomes of provinces in the years 1980 and 1997. Taking spatial dependence into account will give more accurate and informative results about growth process of provinces. Hence, in the first section of this chapter, spatial dependence among the provinces of Turkey will be discussed. In the proceeding sections, convergence in Turkey will be tested using both traditional approach and distribution dynamics approach.

IV.1 Measurement of Spatial Dependence Among Provinces of Turkey

In order to test spatial dependence in the levels and growth rates of per capita incomes of provinces in Turkey, Moran's I statistic discussed in Chapter III is used. Moran's I statistics for the initial and end year values of natural logarithm of per capita GDP and for the annual growth rate of per capita GDP in the period are calculated. For all the variables, Moran's I is positive indicating the possibility of positive spatial dependence. For natural logarithm of per capita GDP in 1987 and in 2001, Moran's I statistics are 0.38 and 0.65, respectively. For the growth rate of per capita income, the statistic is 0.11, lower than the values for per capita GDP in 1987 and 2001.

Jarque-Bera test for normality of these variables reveal no evidence of departures from normality, thus the significance tests of Moran's I are based on normal distribution. For GDP per capita in 1987 and 2001, standardized Moran's I values are 5.1 and 8.6, respectively, revealing strong significance of spatial dependence. On the other hand, standardized Moran's I for growth rate of per capita GDP is 1.54 corresponding to 12 % significance level and spatial independence in growth rate of per capita income cannot be rejected. Consequently, whereas per capita incomes of provinces have positive spatial autocorrelation, positive spatial autocorrelation in growth rates is not significant.

A useful way to visualize the spatial association is to use a Moran scatter plot, which plots the standardized variable of a province against its spatial lag. Figures 2 through 4 show Moran's scatter plots of initial and end year per capita income and growth of per capita income.

In the Figures 2 through 4, points in the quadrant 'HH' show the provinces that have high values themselves and high values in the neighbor provinces, points in 'HL' show the provinces that have high values themselves but low values in the neighbor provinces, points in 'LL' show the provinces that have low values themselves and low values in the neighboring provinces and points in 'LH' show the provinces that have low values themselves but high values in neighbor provinces. Therefore, if the distribution is denser in quadrants 'HH' and 'LL' than in quadrants 'HL' and 'LH', then there is positive spatial autocorrelation, and vice versa. If the densities of quadrants are similar, than it seems to be no autocorrelation in the distribution.

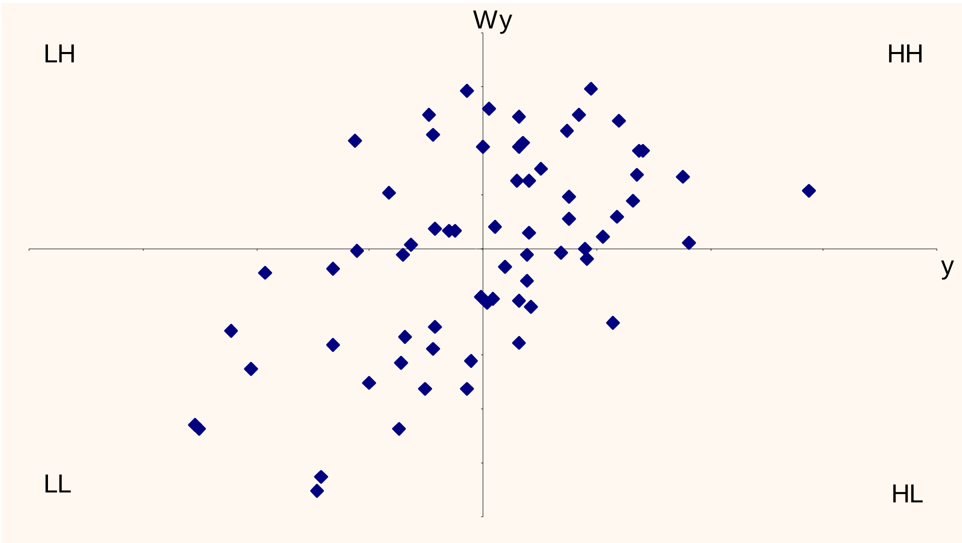


Figure 2: Moran Scatterplot of (log of) GDP per capita (1987)

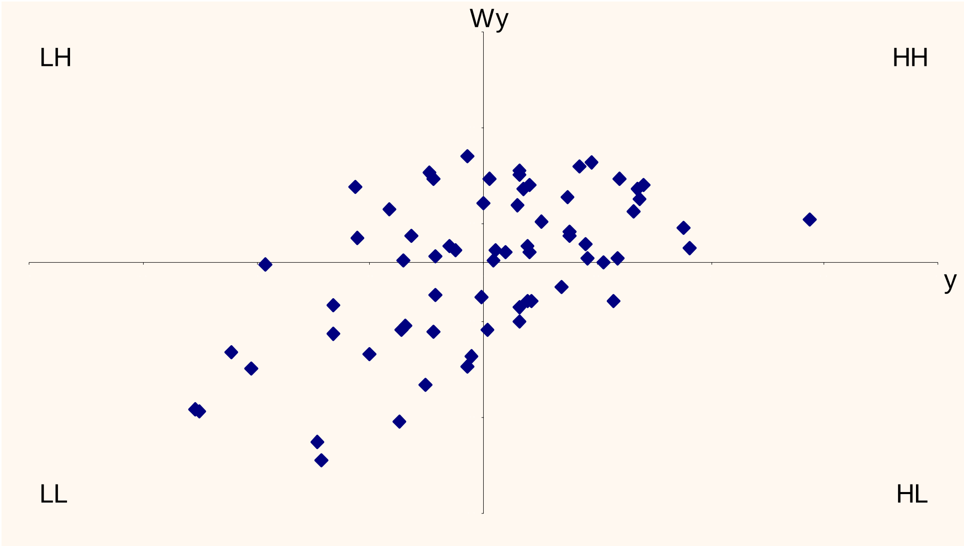


Figure 3: Moran Scatterplot of (log of) GDP per capita (2001)

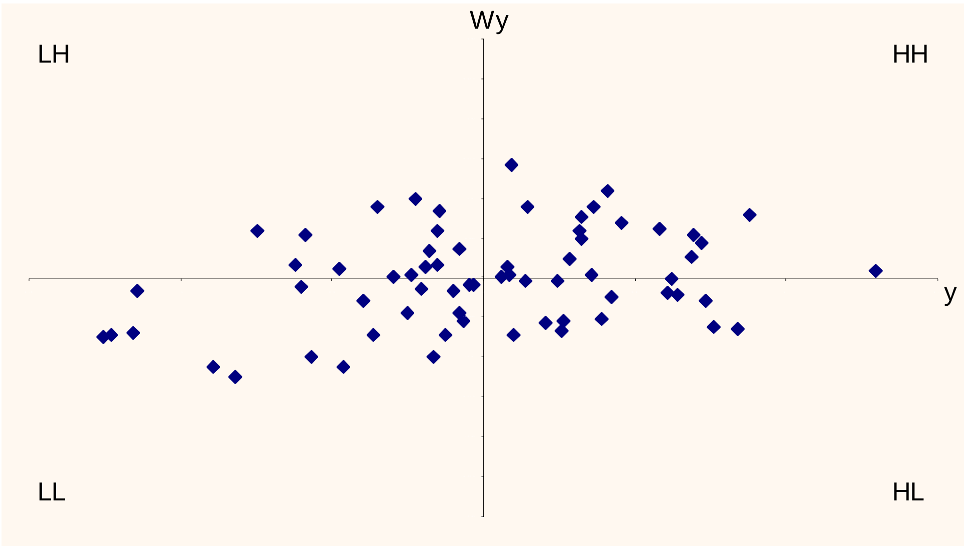


Figure 4: Moran Scatterplot growth rate of per capita GDP (1987-2001)

Moran scatter plots show positive autocorrelation in both levels and growth rates of per capita GDP in the provinces of Turkey supporting the findings from Moran's I statistics. Both in 1987 and in 2001, per capita GDP of 47 of 67 provinces (70%) lie in the quadrants HH and LL. In the growth rate of per capita GDP, the result is weaker. 39 of the 67 provinces (58%) lie in the quadrants 'HH' and 'LL'.

Analysis of spatial dependence reveals that per capita incomes of the provinces in Turkey are spatially autocorrelated. Therefore, in analysis of convergence, spatial dependence must be tested and integrated in the analysis.

IV.2 Traditional Approach to Convergence

In this section, traditional approach discussed in chapter II will be implemented to Turkish data. First, results that do not take spatial dependence into account will be given. Then, spatial dependence will be integrated to the analysis.

IV.2.i Basic Results without Spatial Dependence

As discussed in chapter II, traditional approach uses cross-section regressions in analyzing convergence. If a negative relation between growth rates of economies and their initial income levels are found, then β convergence is said to occur. Another convergence concept, which is used as a complement to β convergence in the traditional approach, is σ convergence. There exists σ convergence if the dispersion in per capita incomes decreases over time.

We start by examining σ convergence across the provinces of Turkey. Coefficient of variation of per capita incomes across provinces around the national per capita income of Turkey is used as a measure of income dispersion. Figure 5 shows the result.

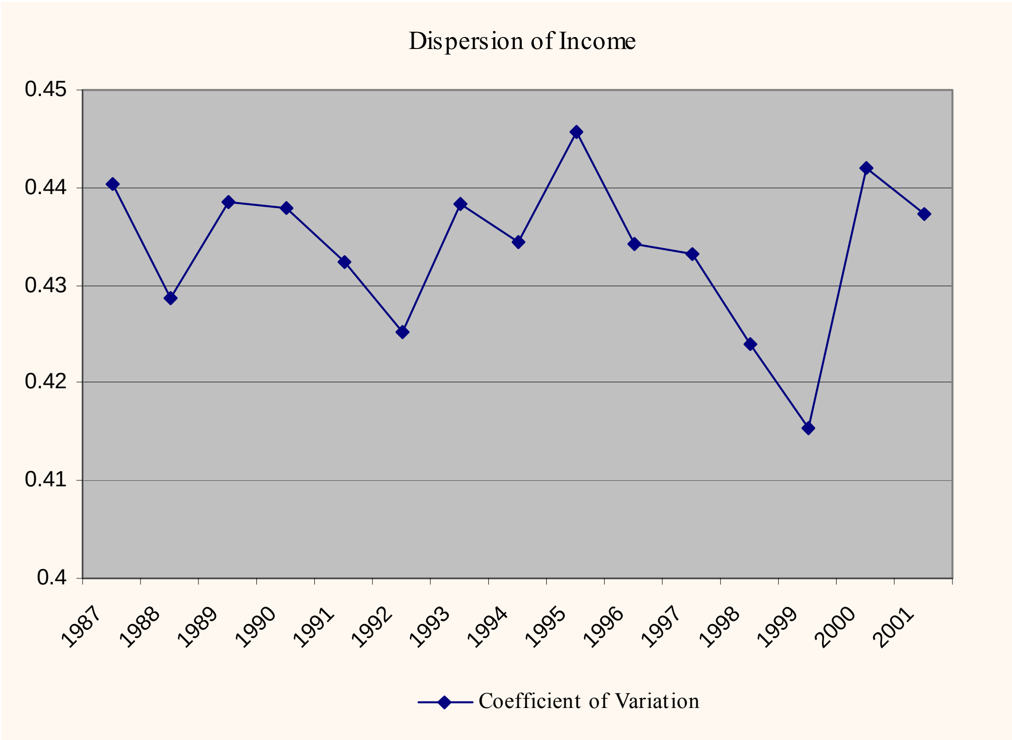


Figure 5: Dispersion of Income in Provinces of Turkey.

As Figure 5 shows, there is not a significant difference in the period of 1987-2001 in coefficient of variation. In 2001, coefficient of variation was about 44% as in 1987 and in the period it fluctuated between 41% and 44%. There is not a downward trend in coefficient of variation. Therefore, there seems to be no convergence in the period.

Analysis of σ convergence concludes that there is no convergence in Turkey. We now turn to β convergence analysis. As discussed in Chapter II, β convergence is a necessary but not a sufficient condition for σ convergence. Therefore, there may be β convergence even if there is no evidence for σ convergence.

In this study, unconditional β convergence equation for the provinces in Turkey is estimated. There are mainly two reasons for preferring unconditional β convergence. First, although differences in technology and preferences do exist across regions within a country, these differences are likely to be smaller than those across countries since regions within a country share a common central government, institutional and legal system. Second and more important, as a policy issue, conditional β convergence is irrelevant. One cannot argue that there is convergence and policies to reduce regional disparities are successful using conditional β convergence framework.

In order to test absolute β convergence, growth rates of provinces are regressed on their initial per capita income levels. The estimated convergence equation is

$$\frac{1}{T}\log\left(\frac{\hat{y}_i}{\hat{y}_{i_0}}\right) = 0.05 - \left(\frac{1-e^{-0.003t}}{T}\right)\log(\hat{y}_{i_0}) \qquad \text{IV.1}$$

where $\hat{y}_i$ is the end year per capita income of province *i* (in 2001*),* $\hat{y}_{i_0}$ is the initial per capita income of province *i* in the initial year (1987) an T is the time period between end year and the initial year (14 years). The rate of convergence is quite low (0.3 %) and insignificant with a t-value of 0,9. The R^2 of the regression is only 0.01. Therefore, we cannot conclude that poor provinces grow faster than the rich ones and that there is convergence in per capita incomes of the provinces in Turkey. The findings are in line with the earlier studies such as Erk *et al*. (2000) and Karaca (2004).

IV.2.ii Integration of Spatial Dependence

Spatial dependence in regional data may cause residuals of the convergence regression to be autocorrelated. As discussed in section 1, per capita incomes of provinces are spatially autocorrelated both in initial and end years. Hence, spatial autocorrelation in the residuals of convergence equation IV.1 must be tested.

In order to check spatial autocorrelation, Moran's I statistic, likelihood ratio (LR) test, Wald test and lagrange multiplier test of OLS residuals (LMERR) are used. The results of these test statistics in the cross-sectional regression of convergence in section 4.1 are given in Table 1.[5]

Table 1: Spatial Autocorrelation in OLS Model.

	Test Statistic	p-value
Moran's I	0.18	0.012
LR	5.67	0.017
Wald	7.75	0.005
LMERR	4.95	0.026

In the β-convergence model, all of the test statistics detect existence of autocorrelation in the least squares estimation. Therefore, specification of unconditional β-convergence equation must be reconsidered to include spatial dependence.

The first model used to take into spatial dependence is the spatial cross-regressive model, which includes the spatial lag of the initial per capita income to the β-convergence model leading to

$$g = \alpha + b\log(\hat{y}_0) + \theta W \log(\hat{y}_0) + u \qquad \text{IV.2}$$

where g is the vector of growth rate of per capita income throughout the period, $\log(\hat{y}_0)$ is the vector of natural logarithm of initial per capita income and $W\log(\hat{y}_0)$ is the spatial lag of natural logarithm of initial per capita income, that is the average of natural logarithm of the initial per capita incomes of the contiguous provinces. The error term u satisfies the Gauss-Markov assumptions. The spatial

[5] Applications of spatial autocorrelation tests and estimations of spatial models are done by spatial econometric toolbox applicable in Matlab by James P. Sage of University of Toledo. Codes can be reached at www.spatial-econometrics.com internet site.

cross-regressive model is estimated via OLS since the spatial lag of per capita income is exogenous.

The results of the cross-regressive model are given in Table 2. The model has interesting findings. First, when the initial per capita levels of neighbor provinces are taken account, the β coefficient is significantly negative that is provinces with lower initial per capita income levels grow faster. This result is conflicting with the results of OLS model with no spatial effects taken into account. However, the speed of convergence is only 0.7% per year and the corresponding half-life is 108 years.[6] Therefore, the rate of convergence is very low compared to the findings of Barro and Sala-i Martin. Second, and more interestingly, the coefficient of spatial lag of initial per capita income is significantly positive. Therefore, provinces with richer neighbors grow faster than the ones with poor provinces.

Table 2: Spatial Cross-Regressive Model Estimates

Variable	Coefficient	t-ratio	p-value
Constant	-0.041	-0.75	0.45
$\ln y_0$	-0.006	-1.98	0.05
$W\ln y_0$	0.010	2.16	0.03

Use of goodness of fit measures may be misleading in spatial econometrics especially when the error term structure is non-spherical. Therefore, an R^2 measure calculated in the usual manner is meaningless and may yield nonsensical values (Anselin 1988). Therefore, information based criteria are used for model comparisons throughout the chapter. OLS model of section 1 with no spatial autocorrelation and the spatial cross-regressive model are compared with respect

[6] The models that take spatial dependence into consideration use the linear specification $g = \alpha + b\log(\hat{y}_0)$ with g denoting the growth rate and $\hat{y}_0$ denoting the initial level of per capita income. The rate of convergence β is calculated as $\beta = -\ln(1+Tb)/T$. The half life, that is the time necessary to fill half of the variation to the steady-state is calculated as $\tau = -\ln 2/\ln(1+b)$.

to Akaike Information Criteria, Schwarz Criteria and Hannan-Quinn information criteria.

Formally the information criteria can be written as,

$$\text{Akaike Information Criterion (AIC)} = \frac{-2l}{n} + \frac{2k}{n} \qquad \text{IV.3}$$

$$\text{Schwarz Criterion (SC)} = -\frac{-2l}{n} + k\frac{\ln(n)}{n} \qquad \text{IV.4}$$

$$\text{Hannah-Quinn Criterion (HQ)} = \frac{-2l}{n} + 2k\frac{\ln(\ln(n))}{n} \qquad \text{IV.5}$$

where k is the number of estimated parameters, n is the number of observations, and l is the value of the log likelihood function and k is the number of estimated parameters. They are all based on the minus two times the value of log likelihood function but are adjusted by different penalty functions. Lower values of information criteria point out better models. The comparison results are given in Table 3.

Table 3: Comparison of OLS and Spatial Cross-Regressive Models.

Model	AIC	SC	HQ
OLS	-5.98	-5.91	-5.95
Spatial Cross-Regressive	-6.02	-5.93	-5.98

Spatial cross-regressive model has slightly smaller AIC, SC and HQ values than the OLS model. Therefore, it seems better than the OLS model. However, the crucial question is whether the cross-regressive model clears out spatial autocorrelation in the residuals. Since spatial cross-regressive model is estimated via OLS, statistics for spatial dependence in the residuals of the OLS can be used for the residuals in the spatial cross-regressive model as well. The results of the tests are given in table 4.

Table 4: Spatial Dependence in Cross-Regressive Model.

	Test Statistic	p-value
Moran's I	0.15	0.03
LR	3.59	0.06
Wald	3.37	0.07
LMERR	3.50	0.06

The results of the test statistics are consistent. Moran's I statistic strongly points out spatial autocorrelation in the residuals of the spatial cross-regressive model. The other three statistics conclude although not that strongly existence of spatial autocorrelation. Therefore, spatial cross-regressive model may not be enough to filter out spatial dependence.

Another model used in case of spatial autocorrelation is the spatial lag model, which includes the spatial lag of growth rate of per capita income as an additional regressor, which can be written as

$$g = \alpha + b\log(\hat{y}_0) + \rho Wg + u \qquad \text{IV.6}$$

where Wg is the average of growth rates of contiguous provinces (spatial lag of growth) and the error terms satisfy the Gauss Markov assumptions. Since the spatial lag of growth is endogenous the model is estimated via maximum likelihood method. The results of the spatial lag model can be seen in Table 5.

Table 5: Spatial Lag Model Estimates.

Variable	Coefficient	t-ratio	p-value
Constant	0.055	1.481	0.13
$\ln y_0$	-0.003	-1.290	0.20
Wg	0.338	2.217	0.03

The results are different in the spatial lag model. The coefficient of initial per capita income is not significant. There is no tendency of convergence after including the spatial lag of growth. However, spatial lag model is also inadequate to filter out spatial dependence since Lagrange Multiplier statistic based on errors of spatial lag model (LMLAG), which investigates the spatial errors in the residuals of the spatial lag model, strongly rejects the null hypothesis of spatial independence in the errors. The value of LMLAG statistic is 192.2 (p-value is zero). Therefore, the spatial lag model cannot eliminate the problem of spatial autocorrelation.

Another model that deals with spatial autocorrelation is the spatial errors model. In this model the error term structure is non-spherical. The model can be written as

$$g = \alpha + b\log(\hat{y}_0) + u$$

$$u = \lambda Wu + \varepsilon \quad \text{and} \quad \varepsilon \sim N(0, \sigma_\varepsilon^2 I_n) \qquad \text{IV.7}$$

Use of OLS will give unbiased but inefficient estimators (Anselin 1988). Therefore, maximum likelihood methods are used to estimate the model. Using the fact that $u = \lambda Wu + \varepsilon$ we can write $u = (1-\lambda W)^{-1}\varepsilon$. Then, the model can be rewritten as

$$g = \alpha + b\log(\hat{y}_0) + (1-\lambda W)^{-1}\varepsilon \qquad \text{IV.8}$$

Therefore, in the spatial errors model, a random shock introduced to a specific region will not only affect the growth rate in that region, but will also affect the growth rates of other states through the spatial transformation $(1-\lambda W)^{-1}$ (Rey and Montouri 1999).

Table 6: Spatial Errors Model Estimates

Variable	Coefficient	t-ratio	p-value
Constant (α)	0.089	2.21	0.02
Lny_0 (β)	-0.005	-1.94	0.05
λ	0.417	2.95	0.00

The estimation results of the spatial errors model are given in Table 6. Estimate of λ is quite significant and positive. Therefore, there is positive spatial autocorrelation in the disturbances of the OLS model and a shock to a specific province will affect the growth rate of all provinces positively. On the other hand, β estimate is negative. After filtering out the spatial dependence in the residuals, provinces converge to a common steady state. However, the coefficient is very low and the corresponding convergence speed is only 0.6% (corresponding half life is 122 years) and the coefficient is not strongly significant. (p value is 0.5233).

In order to compare the cross-regressive model, spatial lag model and the spatial errors model, information based criteria are again used. Table 7 shows the results. For all criteria, spatial errors model has the smallest value. Given that cross-regressive and spatial lag models lack to filter out spatial dependence in errors and spatial errors model has the smallest values in all information criteria, spatial errors model seems to be most appropriate model among the three spatial models investigated.

Table 7: Comparison of Spatial Models.

Model	AIC	SC	HQ
Spatial Cross-Regressive	-6.02	-5.93	-5.98
Spatial Lag	-6.70	-6.60	-6.64
Spatial Errors	-6.72	-6.62	-6.68

To sum up, it is clear that the cross sectional convergence equation suffers from spatial autocorrelation. Alternative models to filter out the spatial autocorrelation are used. Spatial errors model seems to be most appropriate model for spatial dependence. In this model, provinces tend to converge but with a very low speed after filtering out spatial dependence.

IV.3 Distribution Dynamics Approach to Convergence

In this section, convergence in the provinces of Turkey will be analyzed using the distribution dynamics approach discussed in Chapter II. As in traditional approach, first basic results without spatial dependence will be given and then spatial dependence will be integrated to the analysis.

IV.3.i Basic Results without Spatial Dependence

The first task to investigate the distribution dynamics of GDP per capita of provinces in Turkey is to form classes in which per capita income for each province will be placed. In order to form classes, GDP per capita of all provinces are normalized by national average for all years in the period, that is

$$\tilde{y}_{it} = \frac{\hat{y}_{it}}{\bar{\hat{y}}_t} \qquad \text{IV.9}$$

where $\tilde{y}_{it}$ is the nationally normalized per capita income of province *i* in year t, $\hat{y}_{it}$ is the per capita income of province *i* in year t and $\bar{y}_t$ is the per capita income of Turkey in year t.

Forming classes is somewhat arbitrary since there is no commonly accepted definition of being poor or rich within a country. In order to check whether the number of classes affect the results, the analysis is done by dividing the sample into four and five classes. The entire sample (total number of observations is 1005 since there are 67 provinces and 15 years) is divided into four and five income classes with equal frequencies and the values of observations in the boundaries of the quintiles form the gridlines for classes. The bounds of the classes are fixed across the entire period under consideration.

The gridlines for the classes are 51%, 72% and 105% of national per capita income in 4-class Markov chain and 45%, 62%, 81% and 112% of national per capita income in 5-state Markov chain. That is poorest provinces whose GDP per capita are below 51 per cent of national GDP per capita form class 1, provinces with GDP per capita between 51 per cent and 72 per cent form class 2, provinces with GDP per capita between 72 per cent and 105 per cent form class 3 and the richest provinces with GDP per capita higher than 105 per cent form class 4, in the 4-class Markov chain.

After forming classes, transitions of provinces between classes throughout the 14-year transition period are found, the transition probabilities are calculated and the transition probability matrices are formed. Estimated transition probability matrices for 4-class and 5-class Markov chains are given in Tables 8 and 9, respectively.

Table 8: Transition Probability Matrix (4 Classes)

Classes	1	2	3	4	N
1	0.94	0.06	0	0	236
2	0.06	0.90	0.04	0	233
3	0	0.05	0.87	0.09	234
4	0	0	0.08	0.92	235
Initial	0.27	0.22	0.27	0.23	
Ergodic	0.23	0.27	0.24	0.26	

In Tables 8 and 9, classes in the first column denote the initial classes and the classes in the first row denote the final classes after one-year transition period. Last column shows the number of transitions for each class throughout the whole period. The entries inside the tables show the corresponding transition probabilities. For example, in Table 8, there are 236 transitions whose initial class is class 1 and a province initially at class 1 in year *t* will be in class 1 in year *t+1* with a probability of 0.94 and in state 2 with a probability of 0.06.

Table 9: Transition Probability Matrix (5 Classes)

Classes	1	2	3	4	5	N
1	0.93	0.07	0	0	0	189
2	0.07	0.83	0.09	0	0	187
3	0	0.10	0.80	0.10	0	187
4	0	0	0.10	0.84	0.06	188
5	0	0	0	0.05	0.95	187
Initial	0.18	0.21	0.20	0.21	0.19	
Ergodic	0.21	0.20	0.18	0.19	0.22	

The eigenvalues of both matrices are smaller than or equal to 1. Therefore, both matrices are ergodic. The ergodic distributions are also given as well as the initial distributions (distribution in 1987).

The transition probabilities show high degree of persistence especially in the poorest and richest provinces in both transition matrices. The probability of a province initially in state 1 to jump up to a higher state is only 0.06 in Table 8 and 0.07 in table 9, whereas the probability of a province initially in richest class to end up in a poorer class is only 0.08 in Table 8 and 0.05 in table 9. In the middle classes there is more mobility in both upward and downward directions. However, these states are also immobile since the diagonal entries are not less than 0.80. Therefore, there seems to be very low interclass mobility and the probability of poor provinces catching the richer ones and jump up to a richer class is very low.

The degree of mobility of states can also be analyzed using mobility indices, which summarize the information about mobility from the transition matrix into a single statistic. Two mobility indices are used. The first one is the Prais index which is formulated as

$$M_1 = \frac{K - tr(P)}{K - 1} \qquad \text{IV.10}$$

where K is the number of classes and $tr(P)$ denotes the trace of transition matrix P. The second index can be written as

$$M_2 = 1 - |\lambda_2| \qquad \text{IV.11}$$

where $|\lambda_2|$ is the absolute value of the second largest eigenvalue of the transition matrix P. For both statistics, values near 1 reveal high interclass mobility and values near 0 show low interclass mobility. For 4-class Markov chain, the values of M_1 and M_2 are 0.12 and 0.03, respectively, which are close to 0. Similarly, for 5-class Markov chain, the values of M_1 and M_2 are 0.16 and 0.03, respectively. Therefore the finding that there is no interclass mobility is also verified by the mobility indices.

The ergodic distribution can be interpreted as the long-run equilibrium distribution given that there is no policy change or external shock. If there is convergence, the frequencies of middle-income classes - especially class 3 in 4-class Markov chain and class 4 in 5-class Markov chain, which include national average-, should be higher than the frequencies of rich and poor classes in the ergodic distribution. Concentration of the frequencies in two different classes, on the other hand, can be considered as formation of clubs, where two groups of provinces converge within each other but the groups do not converge.

Ergodic distributions of both Markov chains reveal no sign of tendency to converge. There is no tendency of concentration of frequencies in middle-income classes in the ergodic distribution. Indeed, the probabilities of class 3 in 4-class Markov chain and class 4 in 5-class Markov chain decline in ergodic distributions compared with the initial distributions. On the other hand, there seems to be no sign of club convergence since frequencies of all classes are similar. Therefore, the divergent situation will remain in the long run in the absence of policy shocks.

In order to test time homogeneity, the $Q^{(T)}$ statistic derived in chapter III is used. Two $Q^{(T)}$ statistics are calculated. In the first one the whole sample is divided into two sub samples. The first sub sample covers the years between 1987 and 1994 and the second sub sample covers the years between 1994 and 2001. In the second Q statistic, all of the yearly transitions are thought to be different sub samples. Therefore, there are 14 sub periods in the second test statistic. Regarding 4-class Markov chain, the value of the first statistic is 4.5 and the second statistic is 76.7, which are lower than the critical values of chi-squared distribution at 5 per cent significance level with 6 and 78 degrees of freedom, respectively, resulting non-rejection of the null hypothesis of equality of transition probabilities for different periods. Similar results are obtained for 5-class Markov chain. Values of statistics are 7.75 and 100.8, which are lower than the critical values of chi-squared distribution of 8 and 104 degrees of freedom respectively. Therefore, the assumption of time homogeneity cannot be rejected.

To sum up, neither of the Markov chains reveal tendency of provinces to converge. Low mobility in the transition matrices indicates that provinces tend to stay in their initial states. Ergodic distributions also reveal that the divergent situation will continue in the long run and there is no tendency of convergence club formation.

IV.3.ii Integration of Spatial Dependence

Traditional approach shows that spatial autocorrelation exists in convergence process of provinces. Thus, spatial dependence should also be included in Markov chain analysis. In order to check effects of spatial dependence, spatial Markov chain analysis, discussed in Chapter III is used.

Spatial Markov chains for 4-class and 5-class Markov chains are given in Tables 10 and 11, respectively. In the tables, first columns give the classes of spatial lag that is the classes which average per capita income of neighbors of a province belong to. Second columns give the corresponding initial classes and the first rows are the final classes. Total numbers of transitions are given in the last columns of each matrix. The entries in the matrix are the corresponding transition probabilities. For example, figure in the second column and third row, 0.97 gives the probability of a province initially in class 1 to stay in class 1, given that its neighbors are in class 1 on average.

Total number of observations in each class reveals that neighboring provinces tend to have similar per capita incomes. In Table 10, among provinces with poorest neighbors (spatial lag 1) total number of observations initially in class 1 is 148 whereas total number of observations in all other states is 54. Therefore, provinces surrounded by poor regions tend to be poor. The same situation is valid for all classes. For all of the spatial lags, the observations are concentrated on the class of spatial lag. On the other hand, as the difference between initial class and the class of the spatial lag increases, number of observations declines. There is even no observation in classes 1 and 2 with spatial lag of class 1. Table 11 reveals similar results for 5-class Markov chain. There are no instances of a province being in class 4 or 5 with spatial lag 1, being in class 1 with spatial lag 4 and being in class 1 or 2 with spatial lag 5. Therefore, per capita income of a province is affected by its neighbors' per capita incomes.

Transition probabilities in spatial Markov chain significantly differ from the traditional Markov chain in 4-class chain. The probability of a province in class 1 whose neighbors' average per capita income is in class 1 to jump up to a higher income group is 3 percent whereas in the entire sample it is 6 percent and in the sample with spatial lag of 2 it is 10 percent. Conversely, the probability of a province to move down to class 1 declines from 6 percent to 3 percent if the spatial lag is 3 and 4 percent if the spatial lag is 2 and increases to 16 percent if spatial lag is 1. For almost all cases, the probability of moving up increases and the probability of moving down decreases as class of spatial lag increases.

Table 11 gives similar results for 5-class Markov chain. Generally, as class of spatial lag increases the probability of jumping up increases and vice versa. Therefore, there is positive spatial autocorrelation and the evolution of per capita income of a province is affected by its' neighbors per capita incomes.

Spatial Markov chains for both class formations confirm the presence of spatial autocorrelation. To test the hypothesis formally, the $Q^{(R)}$ statistic discussed in Chapter III is used. The values of test statistics for 4-class and 5-class Markov chains are 31.9 and 82.7, respectively. Both values are higher than the critical values of chi-squared distribution with corresponding degrees of freedom of 15 and 25 at 5 percent significance levels. Therefore, the null hypothesis of equality of transition probabilities is rejected and there is spatial dependence in transition probabilities.

Table 10: Spatial Markov Chain (4 states)

Spatial Lag	Class	1	2	3	4	N
1	1	0.97	0.03	0	0	148
	2	0.16	0.81	0.03	0	37
	3	0	0.18	0.73	0.09	11
	4	0	0	0.17	0.83	6
	Class	1	2	3	4	N
2	1	0.90	0.10	0	0	70
	2	0.04	0.91	0.05	0	105
	3	0	0.05	0.92	0.03	65
	4	0	0	0.17	0.83	18
	Class	1	2	3	4	N
3	1	0.83	0.17	0	0	18
	2	0.03	0.92	0.04	0	91
	3	0	0.06	0.83	0.11	99
	4	0	0	0.10	0.90	82
	Class	1	2	3	4	N
4	1	-	-	-	-	0
	2	-	-	-	-	0
	3	0	0	0.90	0.10	59
	4	0	0	0.05	0.95	129

Notes: The first column of the table gives the classes of the spatial lag. The second column gives the initial classes I, the first row gives the final classes and the entries inside give the corresponding probabilities. Finally, the last row column gives the number of transitions. For example, there were 148 instances in which initial class was 1 with spatial lag of 1 (first entry in the last row) and the probability of a province initially at class 1 with spatial lag of 1 is estimated as 0.97 (first entry in the third column).

Table 11: Spatial Markov Chain (5 states)

Spatial Lag	Class	1	2	3	4	5	N
1	1	0.96	0.04	0	0	0	98
	2	0.18	0.82	0	0	0	22
	3	0	0	1	0	0	4
	4	-	-	-	-	-	0
	5	-	-	-	-	-	0
	Class	1	2	3	4	5	N
2	1	0.90	0.10	0.00	0	0	72
	2	0.08	0.86	0.06	0	0	72
	3	0	0.11	0.80	0.09	0	44
	4	0	0	0.13	0.71	0.17	24
	5	0	0	0	0.50	0.50	8
	Class	1	2	3	4	5	N
3	1	0.89	0.11	0	0	0	19
	2	0.08	0.76	0.16	0	0	50
	3	0	0.07	0.84	0.09	0	81
	4	0	0	0.26	0.74	0	35
	5	0	0	0	0	1	14
	Class	1	2	3	4	5	N
4	1	-	-	-	-	-	0
	2	0	0.88	0.12	0	0	43
	3	0	0.13	0.76	0.11	0	55
	4	0	0	0.08	0.85	0.08	90
	5	0	0	0	0.03	0.97	79
	Class	1	2	3	4	5	N
5	1	-	-	-	-	-	0
	2	-	-	-	-	-	0
	3	0	0	0.33	0.67	0	3
	4	0	0	0.02	0.94	0.04	50
	5	0	0	0	0.04	0.96	75

Notes: The first column of the table gives the classes of the spatial lag. The second column gives the initial classes I, the first row gives the final classes and the entries inside give the corresponding probabilities. Finally, the last row column gives the number of transitions. For example, there were 98 instances in which initial class was 1 with spatial lag of 1 (first entry in the last row) and the probability of a province initially at class 1 with spatial lag of 1 is estimated as 0.96 (first entry in the third column).

IV.4 Conclusion

Analysis of convergence across provinces of Turkey without spatial dependence reveals no evidence of convergence in Turkey. Both the traditional approach and distribution dynamics approach have the same results. The results are in line with the results of earlier findings.

The results of convergence analysis without spatial dependence for Turkey are quite different than the results of the studies for other economies. The convergence studies using both traditional approach and distribution dynamics approach generally find evidence of convergence in US. The studies for Europe have conflicting results. Analyses using traditional approach generally find conditional β convergence in European Union and absolute β convergence in separate countries such as Germany, Spain, Italy, France and Greece for different time periods although the rate of convergence differ between studies. On the other hand, studies using distribution dynamics approach reveal no evidence of convergence in European Union.

Earlier studies showed that, taking into spatial dependence may change the results of convergence analysis. Given the fact that, there is significant spatial dependence in per capita incomes of provinces of Turkey led to the use of spatial models to convergence both in traditional and distribution dynamics approach. The results showed that there is positive spatial autocorrelation in the residuals of β convergence equation and in the transition probabilities of the Markov chain. The parameter estimates in the traditional approach and the transition probability estimates in the distribution dynamics approach change. These results are similar with the findings for US and Europe. However, there is no evidence of convergence after filtering out spatial dependence, either.

CHAPTER V

CONCLUSION

This study analyzes regional income convergence in Turkey taking into account spatial dependence. Regional disparities have been one of the most important problems in Turkey, which is also recognized by the policy makers. Policies to reduce regional disparities and enable income convergence have been implemented since 1960's. The issue of convergence will be even more important in the accession of Turkey to the European Union where the projects to reduce regional disparities have an important share in the Union's budget. Therefore, analyzing income convergence is important for checking the success of regional development policies.

This study was different from the previous studies for Turkey in the sense that spatial dependence in economic growth was taken into consideration. Some studies analyze the convergence process in Turkey but they do not take spatial dependence into account. However, recent studies for US and Europe showed that the results of convergence analysis change significantly when spatial dependence is taken into account. On the other hand, Gezici and Hewings (2002) found significant spatial dependence in per capita incomes of provinces in Turkey. Thus, spatial dependence may change the results of convergence analysis in Turkey and it should be taken into account.

Two alternative approaches to test convergence frequently used in the literature, traditional approach and distribution dynamics approach were applied to analyze income convergence across provinces in Turkey. First, the methods were used without taking the effects of spatial dependence into account, as in the classical convergence literature. The results were in line with the results of the earlier studies for Turkey. There is no tendency of convergence across provinces of Turkey and regional disparities tend to remain when spatial dependence is not taken into account.

Second, spatial dependence were integrated into the analysis of convergence. In the traditional approach, it was shown that residuals of cross-section regression to test convergence suffer from spatial dependence. In addition, an external shock to a province will influence its neighbors positively. In the distribution dynamics approach, the analysis of spatial dependence showed that the probability of a province to move up to richer classes increases, as the neighbor provinces get richer. Therefore, spatial dependence affects the convergence process of provinces. However, there is no evidence of convergence even after controlling for spatial dependence.

The finding of spatial dependence among provinces in terms of per capita income has important implications for regional studies. In econometric studies using regional data, one should investigate the existence of spatial dependence of variables under interest. If there is spatial dependence, an appropriate model that filters out spatial dependence should be used. Use of spatial econometrics will avoid autocorrelation in the error terms and thus misleading results.

In sum, the finding of no tendency to converge suggests that regional development policies have not been successful in Turkey. Some new policy measures should be taken. The finding of significant spatial dependence in convergence of provinces suggests that taking spatial dependence into consideration may be useful in constructing regional development programs.

An important issue for further research may be to detect local spatial spillovers. This study showed that there is spatial dependence in the provinces of Turkey, globally. Further research should focus on local spatial spillovers and find which provinces affect each other most positively. This research will give the chance to the policy makers to simulate the effects of regional development programs and regional development funds will be distributed more effectively.

REFERENCES

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Dordrecht, Kluwer Academic Publishers,.

Azariadis, C. and A. Drazen (1990), "Threshold Externalities in Economic Development", *Quarterly Journal of Economics*, 105, 501-526.

Barro, R.J. (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics,* 106, 407-444.

Barro R.J. and X. Sala-i Martin (1991), "Convergence Across States and Regions", *Brooking Papers on Economic Activity*, 1,107-182.

Barro, R.J. and X. Sala-i Martin (1992), "Convergence", *Journal of Political Economy*, 100, 223-252.

Barro, R.J. and X. Sala-i Martin (1995), *Economic Growth*, New York, Mc Graw-Hill.

Baumol, W.J. (1986), "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show", *American Economic Review* 76,1072-1085.

Baumont C., C. Ertur and J. Le Gallo (2002), "The European Regional Convergence Process,1980-1995: Do Spatial Regimes and Spatial Dependence Matter?", *Regional Economics Applications Laboratory (REAL) Working Paper* No. 02-T-10 (forthcoming in *International Regional Science Review* 2005).

Bickenbach, F. and E. Bode (2003), "Evaluating the Markov Property in Studies of Economic Convergence", *International Regional Science Review*, 26, 363-392.

Chatterji, M. (1992), "Convergence Clubs and Endogenous Growth", *Oxford Review of Economic Policy*, 8, 57-69.

Chatterji, M. and J.H.L. Dewhurst (1996), "Convergence Clubs and Relative Economic Performance in Great Britain: 1977-1991", *Regional Studies*, 30, 31-40.

Davies, S. and M. Hallet (2002), "Interactions Between National and Regional Development", *Hamburg Institute of International Economics (HWWA) Discussion Paper* No. 207.

De La Fuente, A. (2002), "Regional Convergence in Spain:1965-95", *Center for Economic Policy Research (CEPR) Discussion Paper* No. 3137.

DPT (2000), *Uzun Vadeli Strateji ve Sekizinci Beş Yıllık Kalkınma Planı*, Ankara, Devlet Planlama Teşkilatı.

Durlauf S.N. and P.A. Johnson (1995), "Multiple Regimes and Cross-Country Growth Behaviour", *Journal of Applied Econometrics*, 10, 365-384.

Erk N., Ş. Ateş and T. Direkçi (2000), "Convergence and Growth within GAP Region (South Eastern Anatolia Project) and Overall Turkey's Regions", International Conference in Economics IV, METU.

Friedman, M. (1992), "Do Old Fallacies Ever Die?", *Journal of Economic Literature*, 30, 2129-2132.

Gezici F. and J.D. Hewings(2001), "Regional Convergence and Economic Performance of Peripheral Areas in Turkey", *REAL Working Paper* No: 01-T-13.

Gezici, F. and J.D. Hewings (2002), "Spatial Analysis of Regional Inequalities in Turkey", *REAL Working Paper* No. 02-T-11.

Johnson, P. (2000), "A Non_parametric Analysis of Income Convergence Across the US States", *Economic Letters*, 69, 219-223.

Karaca O. (2004), " Türkiye'de Bölgeler Arası Gelir Farklılıkları: Yakınsama Var mı?", *Turkish Economic Association Discussion Paper* No. 2004/7.

Kosfeld, R., H. Eckey and C. Dreger (2002), “ Regional Convergence in Unified Germany: Spatial Econometric Perspective” in *Advances in Macroeconometric Modeling, Papers and Pro-ceedings of the 3rd IWH Workshop in Macroeconometrics (eds. C. Dreger and H. Galler),* Nomos, Baden-Baden.

Le Gallo, J. (2004), “Space Time Analysis of GDP Disparities Among European Regions: A Markov Chains Approach”, *International Regional Science Review*, 27,138-163.

LeSage, J.P. (2002), *The Theory and Practice of Spatial Econometrics,* Retrieved June 15, 2005 from http://www.spatial.econometrics.com/html/sbook.pdf.

Lopez-Bazo, E., E. Vaya, A.J. Mora and J. Surinach (1999), “Regional Economic Dynamics and Convergence in the European Union”, *The Annals of Regional Science*, 33, 343-370.

Magrini, S. (2004), “Regional (di)Convergence”, in *Handbook of Regional and Economics Volume 4* (*eds. Vernon Henderson and Jacques-Francois Thisse)*, North Holland, Amsterdam.

Michelis, L., A.P. Papadopoulos and G.T. Papanikos (2004), “Regional Convergence in Greece in the 1980s: an Econometric Investigation”, *Applied Economics*, 36, 881-888.

Neven, D. and C. Gouyette (1995), “Regional Convergence in the European Community”, *Journal of Common Market Studies*, 33, 47-65.

Pekkala S. (2000), “Aggregate Economic Fluctuations and Regional Convergence: The Finnish Case 1988-95”, *Applied Economics*, 32, 211-219.

Quah, D. (1993a), “Galton’s Fallacy and Tests of the Convergence Hypothesis”, *Scandinavian Journal of Economics*, 95, 427-443.

Quah, D.T. (1993b), “Empirical Cross-Section Dynamics in Economic Growth”, *European Economic Review*, 37, 426-434.

Quah, D.T. (1996a), "Regional Convergence Clusters Across Europe", *European Economic Review*, 40, 951-58.

Quah, D.T. (1996b), "Empirics for Economic Growth and Convergence", *European Economic Review*, 40, 1353-1375.

Rey, S.J. (2001), "Spatial Empirics for Economic Growth and Convergence", *Geographical Analysis*, 33, 195-214.

Rey S.J. and B.D. Montuori (1999), "US Regional Income Convergence: A Spatial Econometric Perspective", *Regional Studies*, 33, 143-156.

Romer, D. (2001), *Advanced Macroeconomics (second edition).* Mc-Graw Hill, Boston.

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

Solow, R.M. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70, 65-94.

Tansel A. and N.D. Güngör(1998), "Economic Growth and Convergence: An Application to the Provinces of Turkey, 1975-1995", METU ERC *Working Paper* No. 98/9.

Temel T., A Tansel and P.J. Albersen (1999), "Convergence and spatial patterns in labor productivity: non-parametric estimations for Turkey", *Journal of Regional Analysis and Policy*, 29, 3-19.

Upton, J.G. and B. Fingleton (1985), *Spatial Data Analysis by Example*. Wiley, New York.

APPENDIX

ERGODICITY IN MARKOV CHAINS

Markov chains are used to predict the limiting distribution of per capita incomes as well as to present the transition probabilities between states. In order a Markov chain to have limiting distribution, it must be ergodic. In order to check whether a Markov chain is ergodic, some properties of states need to be defined.

If $\phi_{ij}(l)$ denotes the probability that state (class) i is reached for the first time from state j after l periods, the sum

$$\sum_l \phi_{ii}(l) = \phi_{ii}$$

A.1

represents the probability of eventual return to the original state i.

A state i is called transient if ϕ_{ii}< 1. In this case there is a positive probability 1-ϕ_{ii} that starting from state i a region will not return to the same state in a finite number of time periods. If ϕ_{ii}=1, the class is recurrent and the expectation

$$\mu_{ii} = \sum_l l\phi_{ii}(l) \qquad \text{A.2}$$

is the mean recurrence time for class i. If $\phi_{ii}(l) = 1$, the state is absorbing state for which it is true that $\phi_{ii} = 1$ and $\mu_{ii} = 1$. A state is periodic of period $s > 1$ if

$p_{ii}(l) = 0$ except for $l = s, 2s, \ldots.$, where s is the largest integer with this property. A state which is not periodic is called aperiodic. An aperiodic recurrent state with a finite mean recurrence time is called ergodic.

If S is a closed set, any state k outside the set cannot be reached from any state inside the set ($p_{ik} = 0$ for all $i \in S$). If there exists no closed subset other than the set of all states C, the Markov chain is irreducible. Finally, if all classes of an irreducible Markov chain is ergodic, the chain is also ergodic and there exists a stationary distribution, the limiting distribution of the chain, which is independent of time.

The limiting distribution can be calculated by the matrix equation

$$\pi' = \pi' P \qquad \text{A.3}$$

where P is the transition matrix and $\pi = \{\pi_i\}$ is the limiting distribution of states satisfying the condition

$$\sum_i \pi_i = 1. \qquad \text{A.4}$$

The existence of the stationary distribution can be investigated by the second largest eigenvalue, λ_2, of the transition matrix P. If the absolute value of λ_2 is smaller than 1, the cross sectional distribution converges to a stationary distribution where the conditional probability of occupying in the next period is the same as the unconditional probability.

The asymptotic half-life of the chain, hl, that is the amount of time taken to cover half the distance from the stationary distribution, is given by

$$hl = -\frac{\log 2}{\log|\lambda_2|} \qquad \text{A.5}$$

which ranges from infinity when λ_2 is equal to 1 and the chain does not have stationary distribution and 0 when λ_2 is equal to 0 and the system has already reached its stationary equilibrium.