AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

JRAP (1999) 29, 1, 3-19

Convergence and spatial patterns in labor productivity: nonparametric estimations for Turkey

T. Temel, A. Tansel, and P.J. Albersen*
Centre for World Food Studies and Middle East Technical University

Abstract. This study examines convergence in aggregate labor productivity levels, using 1975-1990 data on the 67 provinces of Turkey. Markov chains (a nonparametric approach) are applied to characterize the long-run tendency of labor productivity. Evidence shows polarization: most provinces tend to move toward a low productivity level, while a few move toward a high productivity level. These two groups form convergence clubs around the upper and lower tails of the distribution. Furthermore, nonparametric regression results, in conformity with the results obtained from the Markov chain model, reveal a persistent spatial pattern in labor productivity: a pattern of high productivity that has lasted from 1975 to 1990, around three highly industrialized provinces.

1. Introduction

Convergence in per capita income toward a steady state growth path across countries or regions has been studied extensively in the past decade (Baumol 1986; De Long 1988; Barro 1991; Barro and Sala-i-Martin 1991 and 1992; Mankiw, Romer, and Weil 1992; Temple 1999). The key question in these studies is whether poor economies eventually will catch up to rich ones in terms of both income levels and income growth. The standard approach has been to regress the average growth rate on the initial income level and on a number of conditioning variables, using either time series and/or cross-section data. A significant and negative coefficient for the initial income level is interpreted as evidence for convergence, verifying the prediction of the neoclassical growth model.

---

* We would like to thank Nazim K. Ekinci and Danny T. Quah for priming us on some of the readings. T. Temel and P.J. Albersen are associated with the Centre for World Food Studies (SOW-VU), Amsterdam, and A. Tansel is with the Middle East Technical University, Department of Economics, Turkey.

Tansel and Güngör (1997) and Filiztekin (1997) also adopt this standard approach to investigate convergence in labor productivity and in per capita income across the provinces of Turkey, respectively. The former examined whether the less developed provinces were converging in labor productivity levels and productivity growth rates toward the richer provinces and found absolute convergence in productivity, with a convergence rate faster in the 1980-1995 period of economic and financial liberalization than in the 1975-1995 period. Convergence rates were faster when somewhat homogeneous relatively poorer eastern and relatively richer western provinces were analyzed separately. This study also found a faster speed of convergence when differences in steady states and human capital were taken into account. Filiztekin used a data set from the same time period (1975-1990) as Tansel and Güngör did and found divergence in per capita income among the provinces of Turkey. The only difference between the two studies is that the latter uses per capita income rather than labor productivity (i.e., income per labor force). This contradictory result implies a negative relationship between population and labor force. In light of migration from rural to urban areas and clustering around three industrial provinces (Istanbul-Izmit, Izmir, and Adana), this negative relation is a reasonable inference.

The current study applies Markov chains to examine convergence while performing a nonparametric regression to detect spatial patterns in labor productivity in Turkey. Using province level labor productivity as units of observations, this study concentrates on the period 1975-1990. The Markov model first derives time-invariant distribution of provincial labor productivity, which further enables us to detect regularities that intradistribution dynamics contain. The main advantage of this model is that it allows us to examine how the top 10 percent of the distribution behaves relative to the bottom 10 percent. Such analysis is not in the domain of parametric investigations of convergence, which basically approximate the average behavior of the observations. The critical point is that this behavior will remain unchanged when the observations at the top and bottom 10 percent interchange their places. This is the case where one needs the Markov model to detect this cross-movement of observations in the top and bottom parts of the distribution. In the context of the Markov model, convergence is said to occur if long-run forecasts of the movements approach zero as the forecast horizon grows. [Quah (1996a, 1996b, 1996c) further examines the technicalities involved in the Markov model.]

Markov chains provide useful representations of dynamic processes; however, they have several shortcomings. First, Markov chains do not explain why provinces experience changes in their labor productivity levels over time. They simply describe the probabilities associated with transitions from one state to another. Second, the procedure is cumbersome when more than one or two additional variables are introduced into the analysis. Third, Markov chains have limited capacity to deal with measurement error. Simple Markov models assume that all observed changes are true changes. But when the variables of interest are survey responses, observed changes are somewhat unreliable. Finally, the time-invariant probabilities depend on an a priori grouping or stratification of the observations.

Evidence in this paper supports polarization of provinces—some provinces tend to have high productivity, while some have low productivity. These two groups form convergence clubs around the upper and lower tails of the time invariant (ergodic) distribution. Initially low (high) productivity provinces are more likely to have low (high) levels of productivity in the longer term. Central to this finding is the presence of a persistent spatial pattern in productivity, indicating concentration around three highly industrialized provinces. Although the techniques applied by the current study are not completely comparable to those used by Filiztekin (1997) and Tansel and Güngör (1997), our results are in conformity with those found by Filiztekin but in contradiction with those found by Tansel and Güngör.

2. Methodology

A Markov chain model, employed in various contexts by Stokey, Lucas, and Prescott (1989), Quah (1993a, 1996a), and others, is applied to trace movements within a distribution. In our context, this model is used to obtain information on four characteristics of the dynamically evolving distribution of provincial labor productivity levels: external shapes, intradistribution dynamics, long-run behavior, and the speed of convergence.

Let F_t denote the distribution of the odds between individual provincial productivity level and Turkey's average labor productivity. Assume that this distribution evolves as

$$F_{t+1} = P' F_t$$

where P is the (n*n) transition probability matrix. The above first-order equation describes the evolution of F_t by mapping F_t into F_{t+1}. An element p_{ij} of P represents the probability that a province in class i in period t will be in class j in period t + 1. Using the minimum variance criterion of Cochran (1966), the distribution F_t is somewhat arbitrarily partitioned into n intervals. According to this criterion, within-class (or interval) variance is minimized on the basis of labor productivity levels.

There are two important assumptions involving this first-order equation.[1] First, we assume that it is a first order process. Specifically, the probability that a province will be in a particular class in period t + 1 depends only on the province's class in period t and not on its class in the previous periods. In our

[1] Testing procedures for these assumptions are discussed in the appendix.

context, this assumption is reasonable because we only have three periods to analyze. Second, we assume that the transition probability matrix is stationary. The s-step-ahead distribution is given by:

$$F_{t+s} = (P')^s F_t.$$

The time-invariant distribution of provincial productivity could be found when $s \to \infty$. The stationarity implies that the probability that a province in class i in period t will be in class j in period t + 1 is constant over time. A maximum likelihood estimate of this probability is given by:

$$p_{ij} = 1/(T-1) \sum_{t=1}^{T-1} (N_{ij}^t / N_i^t)$$

where N_{ij}^t is the number of provinces moving from class i to j in period t; N_i^t is the total number of provinces in class i during period t; and T is the number of time periods.

2.1. Existence and uniqueness of a time-invariant distribution

If the elements p_{ij}^n of the stationary transition matrices converge to some value as $n \to \infty$, then we conclude that there exists a time-invariant probability that the process will be in class j after a large number of transitions and that this distribution is independent of the initial class. Below we provide definitions referred to in the derivation of a time-invariant distribution and state the theorem, adopted from Ross (1985, p. 132-187), that guarantees the existence and uniqueness of it. (See Debreu and Herstein (1953) for the properties of P, and Feller (1950) for further details about Markov chains.)

Definition 1. Class j is said to be accessible from class i if $p_{ij}^n > 0$ for some $n \geq 0$.

Definition 2. Two classes i and j that are accessible to each other are said to communicate.

Definition 3. For any class i we let f_i denote the probability that, starting in class i, the process will ever re-enter class i. Class i is said to be recurrent if $f_i = 1$ and transient if $f_i < 1$. Class i is recurrent if $\sum_{n=1}^{\infty} p_{ij}^n = \infty$ and transient if $\sum_{n=1}^{\infty} p_{ij}^n < \infty$.

Definition 4. A Markov chain is said to be irreducible if there is only one grouping of classes (that is, if all classes communicate with each other).

Theorem. For an irreducible ergodic Markov chain $\lim_{n\to\infty} p_{ij}^n$ exists and is independent of i. Furthermore, letting $\pi_j = \lim_{n\to\infty} p_{ij}^n$, $j \geq 0$, then π_j is the unique nonnegative solution of $\pi_j = \Sigma^{\infty}_{i=0} \pi_i p_{ij}$, $j \geq 0$ and $\Sigma^{\infty}_{j=0} \pi_j = 1$.

Intuitively speaking, this theorem says that if a constant transition matrix P describes a Markov chain process, and if the process is allowed to work for a long period of time, then a time-invariant distribution eventually will be reached. After a long period of time, the proportions in the various categories would be approximately constant and would not depend upon the proportions that were in these categories at an initial time period. Because N provinces are investigated at every period we might expect that ($N*\pi_i$) provinces would be in class i after a long period of time. This does not mean that we should expect ($N*\pi_i$) provinces to settle in class i, but that after a long period of time, ($N*\pi_i$) provinces can be expected to be in class i. In another analysis after more time, the same number ($N*\pi_i$) of provinces, which are probably not all the same ones, also can be expected to be in class i.

2. 2. Nonparametric regression for spatial analysis of productivity

A nonparametric regression method is applied to detect spatial patterns in labor productivity, employing geographical information. The data on longitude (Z_{1i}) and latitude (Z_{2i}) of provinces are used as explanatory variables Z_i = (Z_{1i}, Z_{2i}) and provincial labor productivity (Y_i) as dependent variable. For a finite sample $\{Y_i, Z_i\}^n_{i=1}$ of size n, a Nadaraya-Watson estimate for E(y|z) is calculated as a weighted average of y,

$$y_{i\theta}(z) = \sum_{i=1}^{n} y_i P_{i\theta}(z, z_i),$$

for

$$P_{i\theta}(z, z_i) = \psi((z_i - z)/\theta) / \Psi_{i\theta}(z) \text{ if } \Psi_{i\theta}(z) > 0 \text{ and } 0 \text{ otherwise}$$

where

$$\Psi_{i\theta}(z) = \sum_{i=1}^{n} \psi((z_i - z)/\theta).$$

The weighting function $P_{i\theta}(z, z_i)$ will sum to 1 for all z_i. The density function $\psi(\varepsilon;\theta)$ has its mode at $\varepsilon = 0$ (i.e., if $z_i = z$ for all i). When θ (the window size) goes to zero, its support goes to zero. The heavier weights are given to the observations with the z_i closest to z. The postulated form of the probability function $P_i(z, z_i)$ determines the shape of the regression function, $y_{i\theta}(z)$.

The intuition behind this nonparametric regression is that the observations (the y_i's, with the z_i closest to z) contain more information on E(y|z) than observations far from z. The Mollifier function, $\psi((z_i - z)/\theta)$, is assumed to be normal, N~(z, θ), where θ is a positive scalar bandwidth number or smoothing parameter that determines the weights to be assigned to observations in the neighborhood of z. The choice of smoothing parameter, θ, plays an important role in nonparametric regression estimations, because it affects the magnitude of the weights assigned to observations in the neighborhood of z. For example, if θ is too large, the observations far from z will have a large impact on $y_{i\theta}(z)$.

Although it is common practice to assume an exogenous smoothing parameter, it is important that this parameter depends on the data with a view to reflecting sample size and scale of measurement. We determine the optimal smoothing parameter using least squares cross-validation techniques to determine the optimal bandwidth that gives the best fit of the nonparametrically estimated regression curve to the actual data. [See Silverman (1986), Nadaraya (1989), Härdle (1990), and Keyzer and Sonneveld (1997) for further reading and an application of this method.]

3. Data and variables

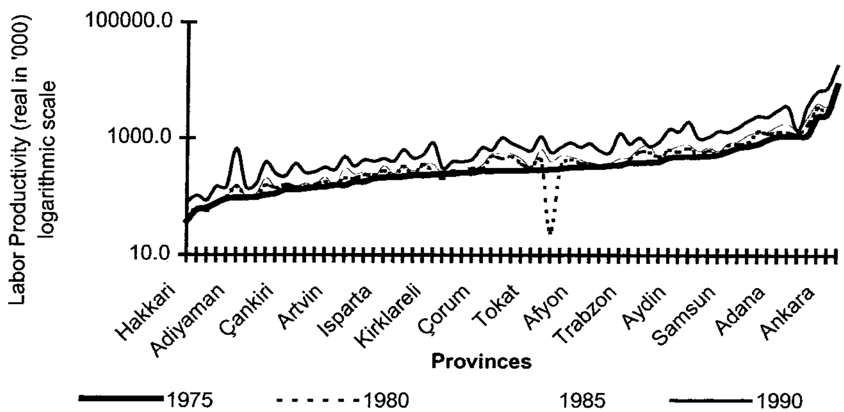
The aggregate labor productivity, which is the gross provincial product (GPP) per worker, for the five-year subperiods between 1975-1990 are used in implementing the above methodology. GPPs for the 67 provinces are taken from Özötün (1980; 1988) for 1975 and 1985 and from the State Institute of Statistics (SIS) (1995) for 1990. The two series from these sources are comparable, although the recent SIS series include new sectors. The data on the employed population (workers) are from SIS (1990). There were 73 provinces in 1990 and 67 provinces in the previous years. GPPs and the number of workers in 1990 were reduced to 67 provinces by adding the figures for the new provinces to their former provinces.

4. Key findings

4.1. Markov chains

The variable of interest F_t is the odds ratio of provincial labor productivity to Turkey's average productivity. To discretize the variable F_{1975}, required by the Markov analysis, we adopt an empirical procedure due to the lack of sound theoretical methods. Our procedure is to first calculate the variable F_{1975} for the initial year 1975 and then sort in ascending order. Next, we divide F_{1975} into intervals in such a way that each interval has minimum variance. [See Cochran (1966) for a detailed exposition of data stratification based on variance minimization.] The jump points in the sorted F_{1975} are accepted as cut-off points for intervals. This procedure helps to determine how much F_{1975} has changed over the period 1980-1990 (Figure 1). The minimum variance criterion suggests six arbitrary intervals: $C_1 = [0, 0.60]$, $C_2 = [0.61, 0.79]$, $C_3 = [0.80, 0.99]$, $C_4 = [1.0, 1.19]$, $C_5 = [1.20, 1.39]$, and $C_6 = [1.40, \infty)$.

Figure 1. Deviations from the 1975 labor productivity level.



The average of P_t over time periods 1975-1980, 1980-1985, and 1985-1990 is used as an estimate of $P = (\Sigma_t P_t/3)$. The elements of the stochastic kernel P (Table 1) are interpreted as follows. The second row indicates that 45 of 268 provinces fall in class 2. Of these, 31 percent moved from class 2 to class 1; 46 percent remained in class 2; 19 percent moved to class 3; and the remaining 4 percent moved to class 4. Furthermore, we observe that those provinces in classes 1 and 6 have high persistence; they tend to remain in the same class as indicated by their respective probabilities of 0.89 and 0.87. Those provinces in classes 2, 3, and 4 are more likely to move to a lower class. The provinces in class 5 are more likely to move to class 6 as off-diagonal elements indicate. The picture that emerges is one where provinces tend toward either low or high productivity classes, with thinning of the middle classes.

The two-period-ahead transition probabilities matrix (Table 1) is used to predict the behavior of provincial productivity levels for the year 2000. The predictions indicate high persistence in classes 1 and 6 and low persistence among the middle classes. Assuming that the same economic structure holds in the future as in 1975-1990, the results suggest an increasing disparity in provincial productivity levels.[2]

[2] Chi-square test results suggest that the transition probability matrix is time-invariant, supporting the observation that during 1975-1990 the economy did not experience significant structural changes.

Table 1. Transition probabilities

Panel A: Transition probabilities matrix (P)

Class	1	2	3	4	5	6	N
1	0.89	0.08	0.02	0.02	0	0	63
2	0.31	0.46	0.19	0.04	0	0	45
3	0.02	0.29	0.47	0.14	0.07	0.02	55
4	0	0	0.23	0.54	0.17	0.06	51
5	0	0	0.11	0.31	0.19	0.39	15
6	0	0	0.03	0.06	0.04	0.87	39
Ergodic distribution	0.35	0.12	0.13	0.13	0.05	0.22	268
Eigenvalue	0.18	0.51	1	0.73	0.92	0.07	

Panel B: Two-period-ahead transition probabilities matrix (P^2)

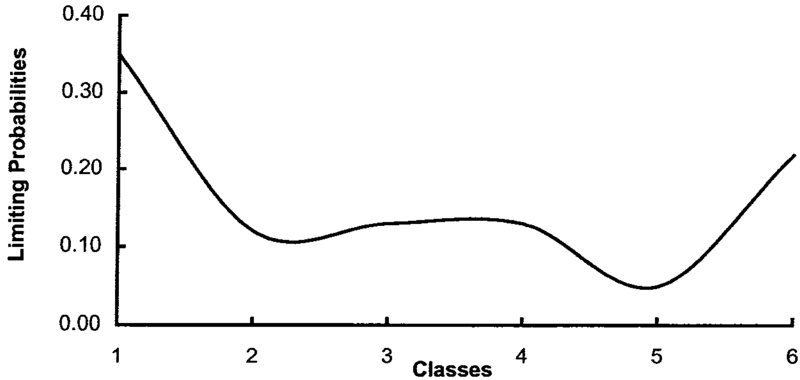
Class	1	2	3	4	5	6
1	0.82	0.11	0.04	0.03	0	0
2	0.43	0.28	0.19	0.07	0.02	0.01
3	0.12	0.27	0.32	0.17	0.06	0.06
4	0	0.07	0.25	0.38	0.14	0.16
5	0	0.03	0.16	0.26	0.11	0.44
6	0	0.01	0.06	0.1	0.05	0.78

Table 1 presents the implied time-invariant (ergodic) distribution[3] of provincial productivity levels (Figure 2), which is the unique solution to the system of equations in our theorem. Everything else constant, time-invariant probabilities indicate that in the long-run the probability of a province staying in classes 1 and 2 is 47 percent and the probability of a province staying in class 6 is 22 percent. The asymptotic behavior of provincial productivity levels imply polarization whereby some of the provinces tend to become poor while others tend to become rich. These two groups of provinces form convergence clubs.[4]

[3] Speed of convergence is the second largest eigenvalue of the kernel P. In our case it is 0.92 which is high. This is the asymptotic rate at which time-invariant distribution is reached. The speed in our context has no relation to the speed implied by the regression analysis. We also calculated two measures of mobility μ_1 and μ_3 as follows: $\mu_1 (P) = (n - tr (P))/(n-1) = 0.52$ and $\mu_3 (P) = 1- |\lambda| = 0.08$. The lower is μ_1, the more persistence there is in the kernel P. The index μ_3 is an asymptotic mobility index that takes high values when P is highly persistent (Quah 1996b).

[4] In the 1975 classification, the following were the provinces starting the process in class 6: Adana, Eskisehir, Bursa, Ankara, Içel, Izmir, Zonguldak, Kocaeli, Istanbul. In 1990 Tekirdag, Bilecik, Kirklareli were added to this list while Zonguldak exited class 6. Similarly, the following provinces were in classes 1 and 2 in 1975: Bingöl, Agri, Hakkari, Adiyaman, Ordu, Gumushane, Kars, Sinop, Van, Bitlis, Yozgat, Erzurum, Tokat, Cankiri, Tunceli, Kastamonu, Mardin, Mus, Giresun, and Erzincan. In 1990, Afyon, Sivas, Trabzon, Nigde, Isparta, and Sanli Urfa were added to this list. Adiyaman exited to class 4, and Corum exited to class 3.

Figure 2. Time-invariant distribution



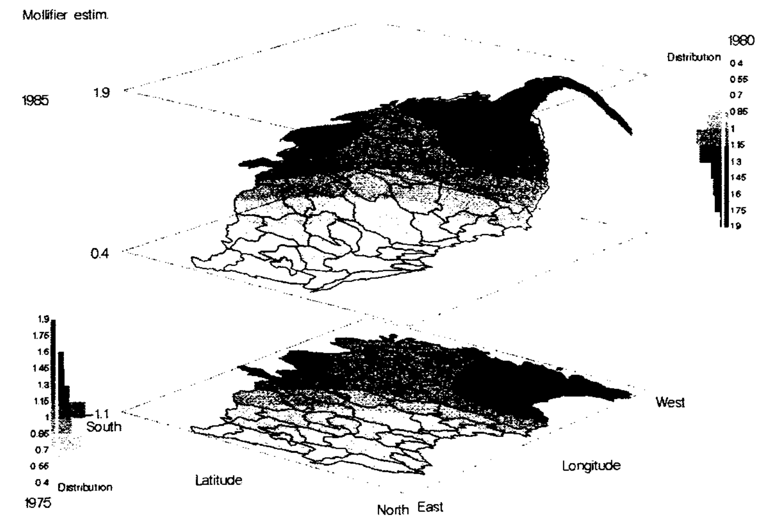
Furthermore, twice as large a probability mass is concentrated in classes 1 and 2 compared to class 6. The vanishing middle classes imply that, in the long run, the majority of provinces have a tendency to move from Turkey's average productivity level, suggesting polarization and divergence. Therefore, productivity levels are not expected to equalize in the long run.

4.2. Spatial pattern in labor productivity

To detect possible spatial patterns in labor productivity, we estimate productivity (y) as a function of latitude (z_1) and longitude (z_2) using a nonparametric regression technique. The intuition behind such a regression is that labor is expected to be more productive if it operates in an enabling environment, one that has enough physical capital, public infrastructure, and regulatory institutions. Such an environment should also attract labor from less developed provinces. Turkey has few provinces endowed with infrastructure and institutions. In addition, because of long-lasting socioeconomic problems that worsened during the last decade, emigration from east and southeast of Turkey toward Adana, Istanbul, and Izmir speeded and further contributed to clustering of population around these provinces.

We present the nonparametric regression results in three-dimensional graphics. Figures 3 and 4 contain condensed information on the developments in three subsequent years. Figure 3 explains four components in detail (as does Figure 4). The legend at the bottom left represents the odds values for 1975; the shading in this legend should be interpreted together with the shading on the plain. For example, the black shade on this legend (the corresponding odds value is 2.3) shows that the area surrounding Istanbul has the highest odds value

Figure 3. Movements in labor productivity, 1975-1985

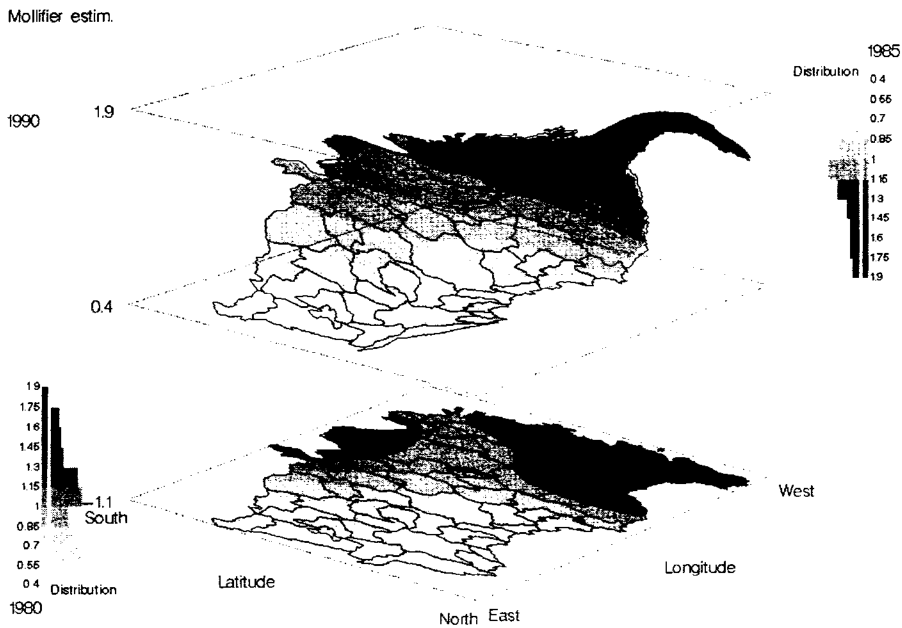


in 1975. The second legend on the top-right corner is associated with the shading of the three-dimensional surface and has the same meaning as that of the bottom left legend. The second legend shows that the area surrounding Istanbul had the highest odds values in 1980 as well. The third component is the shaded plain including the map of Turkey, with which one can locate a province once its longitude and latitude are known. The fourth component is the three-dimensional surface that corresponds to the odds values of the year 1985.

Figure 3 summarizes the movements in labor productivity over the period 1975-1985. The shape of the surface indicates that in 1985 three peaks appeared: one around Adana (the lowest peak), one around Izmir (the middle peak), and one around Istanbul (the highest peak of all). The shading of this surface tells us that the Istanbul region included in 1980 the provinces with the highest productivity. A similar pattern is also present in 1985, suggested by similar shading patterns on the plain and surface. Figure 4 shows that similar developments occurred during 1980-1990 as well, with much stronger concentration around the Istanbul region and the Izmir region.

As shown in Figures 3 and 4, the regression results confirm the results obtained from the Markov chain model. The Markov model predicts in the long

Figure 4. Movements in Labor Productivity, 1985-1990



run a polarization of provinces, while the regression results show that this predicted polarization started in 1975 and became stronger in 1990. A persistent spatial pattern exists in labor productivity over the period 1975-1990.

5. Policy implications and conclusions

This study applies the Markov chain model to the provincial productivity levels across 67 provinces of Turkey from 1975 to 1990. This model is used to project the probability distribution of productivity levels into the year 2000 and to determine the time-invariant distribution and the asymptotic speed of convergence to this distribution.

Assuming that the economic relations over the period 1975-1990 are to prevail in the long run, the calculated time-invariant distribution suggests polarization where most of the provinces move from Turkey's average with accumulations in both low and high tails of the distribution. This observation is also evident in the projections for the year 2000, especially over the long horizon. The grouping of the provinces points to several institutional, technological, and socioeconomic impediments. First, although there is substantial labor mobility across the provinces, there are still large differences in their accumulated human capital because high-income provinces are able to attract highly educated and skilled workers. Such industrialized provinces are placed in the upper tail of the distribution. A recently passed law that extends compulsory education to eight years is a welcome step in the direction of uniform distribution of skilled and educated labor across Turkey in the future. It is well known that capital is reluctant to move to eastern Turkey, in spite of the generous incentives given by the government. As a result, the factor intensities differ between the western, industrialized provinces and the eastern, agricultural provinces. There are also institutional impediments in the financial and other sectors that restrict the diffusion of new technologies. Second, the lack of proper infrastructure and its unequal distribution adversely affect both labor and capital productivity in the eastern provinces, most of which are placed in the lower tail of the distribution.

The nonparametric regression results show that the polarization predicted by the Markov model was present in 1975 and persisted and became stronger from 1980 to 1990. These results further suggest a persistent spatial pattern in labor productivity, in which three major industrialized provinces are the centers of high labor productivity.

A common observation is that a developing country typically has few industrialized provinces, and these provinces naturally become the centers around which population is clustered. This is because urban centers are favored in the allocation of resources and become the areas of high concentration of investment in public infrastructure and public services. This concentration creates an enabling environment for labor in these centers, and enhances its productivity. We perform a spatial analysis to show that such concentration prevails in Turkey. Our spatial analysis suggests that high labor productivity has a public goods feature, meaning that low productivity provinces that are physically close to high productivity provinces slowly benefit from this geographical proximity. A natural extension of this positive externality in our context is that polarization is unavoidable in the early stages of development. But for the development to continue at later stages, policies that mitigate polarization or inequality across provinces should be designed and implemented. In view of this contagiousness of labor productivity, investment in human resources is essential to reduce the existing polarization. Increasing interaction between private enterprises and universities in Turkey shows that investment in human resources is the most effective way to improve the efficiency of labor.

The challenge for future studies is to go deeper and determine the sectoral sources of the polarization in the aggregate labor productivity level found in this study.

References

Barro, R.J., "Economic Growth in a Cross-Section of Countries, " *Quarterly Journal of Economics*, 106, no. 2 (1991), pp. 407-443.

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

Barro, R.J., and X. Sala-i-Martin, "Convergence, " *Journal of Political Economy*, 100, no. 2 (1992), pp. 223-251.

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

Cochran, W., *Sampling Techniques, Second Edition* (New York: John Wiley & Sons, Inc., 1966).

Debreu, G., and I.N. Herstein, "Non-negative Square Matrices, " *Econometrica* (1953), pp. 597-607.

De Long, J.B., "Productivity Growth, Convergence, and Welfare: Comment, " *American Economic Review*, 78 (1988), pp. 1138-1154.

Feller, W., An Introduction to Probability Theory and its Applications, Third Edition, Vol. 1 (New York: John Wiley & Sons, 1950).

Filiztekin, A., "Turkiye'de Iller Arasinda Yakinsama, " ("Convergence Among the Provinces of Turkey"), Koc University Working Paper Series - 1997-15.

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

Goodman, A.L., and W.T. Anderson, "Statistical Inference About Markov Chains, " The Annals of Mathematical Statistics, XXVIII (1957), pp. 89-110.

Goodman, A.L., "Statistical Methods for Analyzing Processes of Change, " *The American Journal of Sociology* (1962), pp. 57-78.

Keyzer, A.M., and B.G.J.S. Sonneveld, "Using the Mollifier Method to Characterize Data Sets and Models: The Case of the Universal Soil Loss Equation, " *International Journal of Aerospace Survey and Earth Sciences* (1997), pp. 263-272.

Mankiw, N., D. Romer, and D.N. Weil, "A Contribution to the Empirics of Economic Growth, " *Quarterly Journal of Economics*, 107 (1992), pp. 407-437.

Nadaraya, E.A. (translated by S. Klotz), *Nonparametric Estimation of Probability Densities and Regression Curves* (Amsterdam: Kluwer, 1989).

Özötün, E., *Iller Itibariyle Türkiye Gayri Safi Yurtiçi HasilasiKkaynak ve Yöntemler 1975-1978 (Provincial Distribution of the Gross Domestic Product of Turkey—Sources and Methods 1975-1978)*, Publication No. 907 (Ankara: State Institute of Statistics, 1980).

Özötün, E., *Türkiye Gayri Safi Yurtiçi Hasilasinin Iller Itibariyle Dagilimi 1979-1986 (Provincial Distribution of the Gross Domestic Product of Turkey 1979-1986)*, Publication No. 1988/8 (Istanbul: Istanbul Chamber of Industry Research Department, 1988).

Quah, D., "Empirical Cross-Section Dynamics in Economic Growth, " *European Economic Review*, 37, no. 203 (1993a), pp. 426-34.

Quah, D., "Galton's Fallacy and Tests of the Convergence Hypothesis, " *Scandinavian Journal of Economics*, 95 (1993b), pp. 427-443.

Quah, D., "Aggregate and Regional Disaggregate Fluctuations, " *Empirical Economics*, 21 (1996a), pp. 137-59.

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

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

Ross, M.S., *Introduction to Probability Models* (New York: Academic Press, 1985).

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

State Institute of Statistics (SIS), data diskettes for provincial gross domestic product (Ankara: SIS, 1995).

State Institute of Statistics (SIS) *Census of Population, Social and Economic Characteristics of Population*, Publication No. 1369, etc. (Ankara: SIS, 1990 and various census years).

Stokey, N.L., R.E. Lucas, and E.C. Prescott, *Recursive Methods in Economic Dynamics* (Cambridge, MA: Harvard University Press, 1989).

Tansel, A., and N.D. Güngör, "Income and Growth Convergence: An Application to the Provinces of Turkey, " paper presented at the First Annual ERC/METU Conference on Economics (September 18-20, 1997).

Temple, J., The New Growth Evidence, " *Journal of Economic Literature*, XXXVII, no. 1 (1999), pp. 112-156.

Appendix

In this appendix we first explain how to test for the two assumptions of a Markov chain: time-stationarity of the transition probability matrices and the first-order Markov property. [See Goodman and Anderson (1957) and Goodman (1962) for a detailed discussion of the test procedures applied in the present paper.] A theoretical framework is provided for the existence of a time-invariant distribution to which the process converges.

For illustrative purposes, the following contingency table will be referred to throughout the appendix:

$$A_t = \begin{array}{cccc} \text{Classes} & 1(t) & 2(t) & \text{Total} \\ 1(t-1) & n_{11}^t & n_{12}^t & n_{1.}^t \\ 2(t-1) & n_{21}^t & n_{22}^t & n_{2.}^t \\ \text{Total} & n_{.1}^t & n_{.2}^t & n^t \end{array}$$

where t = 0, 1, 2, 3 and i = j = 1, 2. Using A_1, A_2, and A_3 and the definitions given in the text, we construct a table with (T*m) (or 3 ×y 2) dimensions:

$$Z_i = \begin{array}{ccc} t/j & j=1 & j=2 \\ t=1 & \hat{p}_{i1}^1 & \hat{p}_{i2}^1 \\ t=2 & \hat{p}_{i1}^2 & \hat{p}_{i2}^2 \\ T=3 & \hat{p}_{i1}^3 & \hat{p}_{i2}^3 \end{array}$$

Assumption 1. The transition probabilities are constant over time. Here the null hypothesis is $H_0 : p_{ij}^t = \hat{p}_{ij}$ for all t. An alternative to this assumption is that the transition probability depends on t, $H_1 : p_{ij}^t = \hat{p}_{ij}^t$ where

$$\hat{p}_{ij}^t = \left(\frac{n_{ij}^t}{n_i^{t-1}} \right)$$

is the estimate of the transition probability for time t. Under these hypotheses, the likelihood ratio is of the form,

$$\lambda = \prod_t \prod_{i,j} \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}^t} \right]^{n_{ij}^t},$$

where

$$\prod_{t=1}^{T} \prod_{i,j} \hat{p}_{ij}^{n_{ij}^t} \text{ hold under } H_0 \text{ and}$$

$$\prod_{t=1}^{T} \prod_{i,j} (\hat{p}_{ij}^t)^{n_{ij}^t} \text{ holds under } H_1.$$

$-2\log\lambda$ is distributed as $\chi^2_{(T-1)[m(m-1)]}$ when H_0 is true. The likelihood ratio resembles likelihood ratios obtained for standard tests of homogeneity in contingency table A_t. The null hypothesis states that the random variables represented by the T rows in Z_i have the same distribution. In order to test it, we calculate

$$\chi_i^2 = \sum_{i,j} n_i^{t-1} (\hat{p}_{ij}^t - \hat{p}_{ij})^2 / \hat{p}_{ij}.$$

If H_0 is true, χ_i^2 has the limiting distribution with (m - 1)(T - 1) degrees of freedom, the set of χ_i^2 's is asymptotically independent, and the sum

$$\chi^2 = \sum_{i=1}^{2} \chi_i^2 = \sum_i \sum_{t,j} n_i^{t-1} (\hat{p}_{ij}^t - \hat{p}_{ij})^2 / \hat{p}_{ij}$$

has the usual limiting distribution with (T-1)[m(m - 1)] degrees of freedom.

Another way of testing the same hypothesis is to calculate

$$\lambda_i = \prod_{t,j} \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}^t} \right]^{n_{ij}^t}$$

for i = 1, 2 using Z. The asymptotic distribution of $-2\log\lambda_i$ is χ_i^2 with (m - 1)(T - 1) degrees of freedom. The test criterion based on λ can be written as

$$\sum_{i=1}^{m} -2\log\lambda_i = -2\log\lambda.$$

Assumption 2. The Markov chain is of a given order. Intuitively speaking, this assumption states that the location of a province at time (t + 1) is independent of its location at time t. A Markov chain is second-order if a province is in class i at time (t - 2), in j at time (t - 1), and in k at time t. Let p_{ijk}^t denote the probability that a province follows a second-order chain. Time stationarity then implies $p_{ijk}^t = p_{ijk}$ for all t = 2, ... , T. A first-order stationary chain is a special case of second-order chain, one for which p_{ijk}^t does not depend on i. Now let n_{ijk}^t be the number of provinces in class i at (t - 2), in class j at (t - 1), and in class k at t. Let

$$n_{ij}^{t-1} = \sum_k n_{ijk}^t \text{ and } n_{ijk} = \sum_{t=2}^{T} n_{ijk}^t.$$

The maximum likelihood estimate of p_{ijk} for stationary chains is

$$\hat{p}_{ijk} = \left(\frac{n_{ijk}}{\sum_{l=1}^{m} n_{ijl}} \right) = \left(\frac{\sum_{t=2}^{T} n_{ijk}^t}{\sum_{t=2}^{T} n_{ij}^{t-1}} \right).$$

The null hypothesis in this case is

$$H_0\text{: } p_{1jk} = p_{2jk} = \ldots = p_{mjk} = p_{jk} \text{ for } j, k = 1, \ldots, m.$$

The likelihood ratio test criterion is

$$\lambda = \prod_{i,j,k=1}^{m} \left[\frac{\hat{p}_{jk}}{\hat{p}_{ijk}} \right]^{n_{ijk}} \quad \text{where } \hat{p}_{jk} = \left(\frac{\sum_{i=1}^{m} n_{ijk}}{\sum_{i=1}^{m} \sum_{l=1}^{m} n_{ijl}} \right) = \left(\frac{\sum_{t=2}^{T} n_{jk}^{t}}{\sum_{t=1}^{T-1} n_{j}^{t}} \right)$$

is the maximum likelihood estimate of p_{jk}. Under the null hypothesis, $-2\log\lambda$ has an asymptotic

$$-\chi^2_{m(m-1)^2}$$

distribution where

$$\chi_j^2 = \sum_{i,k} n_{ij}^{*} (\hat{p}_{ijk} - \hat{p}_{jk})^2 / \hat{p}_{jk} \quad \text{and} \quad n_{ij}^{*} = \sum_{k} n_{ijk} = \sum_{k} \sum_{t=2}^{T} n_{ijk}^{t} = \sum_{t=2}^{T} n_{ij}^{t-1} = \sum_{t=1}^{T-1} n_{ij}^{t}$$

with $(m-1)^2$ degrees of freedom. The corresponding test using the likelihood ratio is

$$\lambda_j = \prod_{i,k=1}^{m} \left[\frac{\hat{p}_{jk}}{\hat{p}_{ijk}} \right]^{n_{ijk}}$$

The asymptotic distribution of $-2\log\lambda$ is chi-square with $(m-1)^2$ degrees of freedom.

To test the joint hypothesis H_0: $p_{ijk} = p_{jk}$ for all i, j, k = 1, 2, … , m, we calculate

$$\chi^2 = \sum_{j=1}^{m} \chi_j^2 = \sum_{j,i,k} n_{ij}^{*} (\hat{p}_{ijk} - \hat{p}_{jk})^2 / \hat{p}_{jk}$$

which has the usual limiting distribution with $m(m-1)^2$. Similarly, the joint test criterion is

$$\sum_{j=1}^{m} -2\log\lambda_j = -2\log\lambda = 2\sum_{i,j,k} n_{ijk} [\log \hat{p}_{ijk} - \log \hat{p}_{jk}].$$