On the Robustness of the Twin-peaked Ergodic Distribution of Income Across Countries

Gisèle Hites*
ECARES, Université Libre de Bruxelles

Abstract

In the literature on convergence, the simple Markov chain model indicates evolution towards a twin-peaked world. Although cleansing the ergodic distribution of income across countries of short-run noise reinforces its twin-peaked shape, these twin peaks are not statistically significant.

Keywords: Ergodic distribution; Filters; Income distribution; Markov chains; Twin peaks.

JEL classification: C23; O57.

* Address: ECARE-CP 139, ULB, 50 Av Roosevelt, 1050 Brussels, Belgium. Tel: 32-2-650-3860. Fax: 32-2-650-4012. E-mail: ghites@ulb.ac.be.

1. Introduction

The simple Markov chain model has been widely used in the social sciences to study the phenomenon of mobility. Empirical applications have included geographic, labor, and social mobility. The prime attraction in this approach lies in the simplicity of the characterization of the steady state. One active area of research in which the long run properties of a panel of data are of particular interest is the study of the distribution of income across countries. In this body of literature, Quah's (1993a,b) application of the simple Markov chain model reveals evidence of a world in which the rich and the poor are diverging to form 'twin peaks'. In what follows, the robustness of this conclusion is assessed. It is shown that although cleansing the ergodic distribution of income across countries of short-run noise reinforces its twin peaked shape, these twin peaks are not statistically significant. Moreover, the specific type of high immobility reflected by the data on income renders the estimated transition matrix particularly prone to the generation of twin-peaked ergodic distributions.

2. Filter

First, let us establish notation. There are a finite number of states *m* $(i = 1,...,m)$ and transitions between these states are observed at regular intervals for a finite length of time T $(t = 1,...,T)$. Let $N(t)$ be the matrix of observed transitions at time *t* where the *ij*th element is $n_{ij}(t)$ (the number of transitions from state *i* to state *j* observed at time *t*), and let $n(t)$ be the distribution of the observations across the states at time *t* where the *i*th element is $n_i(t)$ (the number of observations in state *i* at time *t*). We assume that the observed transitions are generated by a simple (i.e. of order one) time-homogenous Markov chain according to the matrix of transition probabilities *P* where the *ij*th element is p_{ij} (the probability of transiting from state *i* to state *j*). The maximum likelihood estimator of *P* is denoted $\hat{P}$, where the *ij*th element is $\hat{p}_{ij} = \sum_{t=1}^{T} n_{ij}(t) / \sum_{j=1}^{m} \sum_{t=1}^{T} n_{ij}(t)$. The model can then be summarized by the following expression: $n(t+1) = n(t) \cdot P = n(0) \cdot P^{t+1}$. In this paper, we are interested in the long-run tendencies of the distribution of the observations, i.e. in the ergodic distribution $n(\infty) = n(\infty) \cdot P$.

Second, let us review Quah's application of the simple Markov chain model. The data used is the Lapeyres index of annual real per capita income from the Summers and Heston (1991) Penn World Tables for 118 countries (relative to the world average) for the 1962-84 time period. Five possible states are defined by discretizing the set of possible values of relative incomes into intervals at 1/4, 1/2, 1, and 2. The results are presented below.

$$\hat{P}_Q = \begin{array}{ccccc} 0.97 & 0.03 & 0 & 0 & 0 \\ 0.05 & 0.92 & 0.04 & 0 & 0 \\ 0 & 0.04 & 0.92 & 0.04 & 0 \\ 0 & 0 & 0.04 & 0.94 & 0.02 \\ 0 & 0 & 0 & 0.01 & 0.99 \end{array}$$
$$n_Q(\infty) = \begin{array}{ccccc} 0.24 & 0.18 & 0.16 & 0.16 & 0.27 \end{array}$$

(1)

The estimated ergodic distribution clearly indicates an evolution towards a bipolar world of haves and have-nots.

Finally, let us turn to the issue of filtering. In the simple Markov chain model, the estimated transition probability matrix is used to extract information concerning the mobility of countries within the distribution of incomes. This information is camouflaged by two sources of noise. The first is generated by inaccuracies in the data and will not be discussed here. The second results from using continuous data to estimate a categorical model (i.e. from translating continuous data into categorical data by defining income class frontiers) and will be discussed below. Whereas the first source of noise affects the inference in ways that are unknown to us, the second source of noise can be observed directly.

In the simple Markov chain model, transitions represent mobility. In reality, however, transitions can occur for two reasons. Transitions can result from higher (or lower) than average world growth in a country. This is what we call mobility and this is what we would like to measure. Transitions can also result from business cycle type variations in a country's income when the level of income is situated very close to one defining a frontier between classes. This is clearly not mobility and since such transitions are included in our calculation of mobility, it is necessary to purge the data of such noise. In sum, we need to correct for the bias that short run fluctuations in income introduce into the calculation of long run tendencies in the distribution of world incomes.[1]

In order to cleanse the data, it is necessary to tighten the conditions under which a transition is considered to represent mobility. Here this is achieved by requiring a transition to last a minimum number of periods in order for it to be counted as mobility, this minimum number of periods being defined as just over the average number of periods spanned by a business cycle. Four increasingly fine filters are applied to the original data.[2] The first

---

[1] This is an estimation problem arising from the bias introduced into the transition matrix by fitting a categorical to a finite sample of continuous data. In theory, the ergodic distribution is independent of short run noise.

[2] The data used is the same as that used by Quah (1993a,b) with the following adjustments: the data comes from a more recent version of the Penn World Tables (i.e. Mark 5.6). and the sample is composed of the 111 countries for which there is

filter counts transitions as mobility if they last for at least one year, that is if no transition is observed during the year following the initial transition. The second, third and fourth filters do the same for two, three and four year spans, respectively. The results obtained from application of these four filters are presented below:

$$n_{H1}(\infty) = 0.4221 \quad 0.1650 \quad 0.1449 \quad 0.0788 \quad 0.1893$$
$$n_{H2}(\infty) = 0.4079 \quad 0.1240 \quad 0.1226 \quad 0.0893 \quad 0.2562$$
$$n_{H3}(\infty) = 0.5192 \quad 0.1274 \quad 0.0983 \quad 0.0659 \quad 0.1892$$
$$n_{H4}(\infty) = 0.5098 \quad 0.1355 \quad 0.0894 \quad 0.0685 \quad 0.1967$$

(2)

Two observations are interesting to note. First, of the 127 transitions initially counted, almost half are due to short-term fluctuations in income and not the long-term tendencies that we are trying to measure. Indeed, of the 127 transitions initially counted, 29 (15, 12 and 4, respectively) lasted less than two (three, four and five, respectively) years. Second, the application of increasingly fine filters reinforces the twin peaked shape of the ergodic distribution. In very rough terms, comparing the ergodic distribution calculated from the most finely filtered data to the ergodic distribution calculated from the unfiltered data, the part of the distribution falling into the poorest class (left hand peak) increases by 20% to reach 50%, the part of the distribution falling into the richest class (right hand peak) remains at 20%, and the part of the distribution falling into the middle classes decreases uniformly.

In sum, when the data is purged of short run noise, the evidence for twin peaks is reinforced. But just how robust is this evidence? In the next section, statistical inference is carried out.

3. Statistical inference

Anderson and Goodman (1957) study the asymptotic properties of first-order Markov chains and show that, for each state i, under the null hypothesis $\hat{p}_{ij} = \tilde{p}_{ij} \ \ (i, j = 1,...,m)$:

$$\sum_{i=1}^{m}\sum_{j=1}^{m} n_i^* \frac{(\hat{p}_{ij} - \tilde{p}_{ij})^2}{\tilde{p}_{ij}} \sim c^2(m(m-1))$$

(3)

where $n_i^* = \sum_{t=0}^{T} n_i(t)$ and $\tilde{p}_{ij}$ are the transition probabilities under the null hypothesis. However, this result is developed under the assumption that every

---

continuously available data for the period 1960-89, that is the 118 countries used by Quah minus Afghanistan, Sudan, Ethiopia, Liberia, Nepal, Iraq, and Tanzania.

$p_{ij} > 0$ and this is clearly not the case here.[3] Nevertheless, we can circumvent this problem by simply neglecting the zeros in each row. Indeed, we are not interested in testing whether these elements are truly zero; rather, we are interested in testing whether the positive elements are truly the estimated values. This is equivalent to treating the first and fifth rows of the estimated transition matrix as the results from the estimation of a two-state Markov chain model (i.e. $m_1 = m_5 = 2$), and the second to fourth rows as the results from the estimation of a three-state Markov chain (i.e. $m_2 = m_3 = m_4 = 3$). Each row has a c^2 distribution with $m_i - 1$ degrees of freedom and since these rows are asymptotically independent, they can be added to obtain a c^2 distribution with $m(m_i - 1)$ degrees of freedom. So, *m* varies from row to row, and Equation becomes:

$$\sum_{i=1}^{m} \sum_{j=1}^{m_i} n_i^* \frac{(\hat{p}_{ij} - \tilde{p}_{ij})^2}{\tilde{p}_{ij}} \sim c^2 (\sum_{i=1}^{m} m_i)$$

(4)

Using the test presented above, different matrices generating different ergodic distributions were tested under the null. These matrices were created by taking the estimated transition matrix resulting from application of the fourth filter and perturbing the off-diagonal elements in a strategic manner. The following three hypotheses were amongst those not rejected. In the first case, decreasing $\hat{p}_{45}$ by 1%, causes the ergodic distribution to become strongly unimodal in state 1:

$$n_1(\infty) = 0.6188 \quad 0.1640 \quad 0.1078 \quad 0.0825 \quad 0.0269 \tag{5}$$

In the second case, increasing $\hat{p}_{12}$ by 1% and decreasing $\hat{p}_{21}$ by 2%, causes the ergodic distribution to become bimodal in states 2 and 5:

$$n_2(\infty) = 0.1945 \quad 0.2235 \quad 0.1474 \quad 0.1132 \quad 0.3215 \tag{6}$$

In the third case, increasing $\hat{p}_{12}$ by 1%, decreasing $\hat{p}_{21}$ by 2%, and decreasing $\hat{p}_{45}$ by 1% causes the ergodic distribution to become unimodal in state 2:

$$n_3(\infty) = 0.2722 \quad 0.3126 \quad 0.2059 \quad 0.1577 \quad 0.0517 \tag{7}$$

These results show that the confidence region for the estimated transition probabilities is big enough to contain matrices generating a whole range of very different ergodic distributions.

Notice that the nonrobustness of the twin peaks result does not originate in the particularly large size of the confidence region, but rather in

[3] Anderson and Goodman's test statistic is an asymptotic result, valid only for sufficiently large samples (i.e. those for which the expected frequency in each of the cells of the transition matrix is superior to 5).

the extreme sensitivity of the ergodic distribution to perturbations to the estimated transition matrix. Analysis of the elasticities of the elements of the ergodic distribution with respect to the elements of a triple diagonal transition matrix provides an explanation of this fragility.[4] First, the elasticities of the elements of the ergodic distribution with respect to the transition probabilities tend to increase as the off-diagonal elements of the triple diagonal transition matrix decrease. Second, given a 'balanced' triple diagonal transition matrix (i.e. one generating a uniform ergodic distribution), the elasticities of states 1 and 5 of the ergodic distribution with respect to the transition probabilities are higher than those of states 2, 3 and 4 of the ergodic distribution. Intuitively, the relatively high elasticity of the endpoints of the ergodic distribution can be explained by the fact that whereas the three middle states all have two exits (i.e. one to a lower state and one to a higher state), the two endpoints only have one exit. So, whereas blocking one of these single exits necessarily leads to a pile up in one of the endpoints in the long run, blocking one of the other exits leads to a pile up that could potentially be distributed amongst all the states behind the blocked exit via the unblocked exit. So, whenever the estimated transition matrix presents a triple diagonal structure, it is easier to generate modes in states 1 and 5 than in any of the other states, especially if the off-diagonal elements of the transition matrix are particularly small.

5. Concluding remarks

Two lessons emerge from this paper. First, when Markov chain models are fitted to continuous data, a bias towards excess mobility is introduced into the estimated transition matrix. Second, when the estimated transition matrix of a simple Markov chain model presents a certain type of high immobility, the corresponding ergodic distribution is characterized by an extreme fragility of a very particular sort.

Acknowledgements

I would like to thank Marco Lippi and Lucrezia Reichlin for helpful comments on previous versions of this paper.

References

Anderson, T.W., and L.A. Goodman, 1957, Statistical inference about Markov chains, Annals of Mathematical Statistics 28, 89-110.

[4] For the general expressions of the elasticities of the elements of the ergodic distribution with respect to the transitions probabilities along with their analytical derivation, please contact the author.

Quah, D., 1993a, Empirical cross-section dynamics in economic growth, European Economic Review 37, 426-434.
Quah, D., 1993b, Galton's fallacy and tests of the convergence hypothesis, Scandinavian Journal of Economics 95, 427-443.
Quah, D., 1996, Empirics for economic growth and convergence, European Economic Review 40, 1353-1375.
Summers, R. and A. Heston, 1991, The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988, The Quarterly Journal of Economics 106, 327-368.