Using a Markov Chain to Look at Academic Performance at the University Level

Toni Mora  
School of Economics and Social Sciences, Universitat Internacional de Catalunya

Josep-Oriol Escardíbul  
Department of Political Economy and Public Finance, University of Barcelona

Correspondence to: Toni Mora, School of Economics and Social Sciences, Universitat Internacional de Catalunya, Immaculada, 22, 08017, Barcelona (Spain) Phone 0034 932541800 (4511) Fax 0034 932541850. Email: tmora@cir.uic.es

Abstract

Using a Markov chain approach in which the transition matrix contains the overall transitions in grades from high school to university, we compute the expected changes in undergraduate performance between the beginning and the end of a student’s degree. The data are compiled from graduates at the University of Barcelona (Spain) for the period 1996–2003. The empirical results show small changes in the ergodic solution if transition probability of improved grades is increased, at least for the period considered. However, educational policies aimed at increasing transitions from lower to upper states seem to be more effective at lower levels. Likewise, the transitions from high school grades to university grades differ according to the kind of high school that students attended (public, private religious, or private non-religious).

Keywords: Markov chains, quality of education, high school transition to university

Introduction

It is generally agreed that countries should improve their human capital by considering both the level (quantity) and the profitability (quality) of the education they provide. Our analysis focuses on educational quality, since it examines students’ grades as they transition between their high school and the university. Hanushek (1979, 2003) points out that more attention should be given to the distribution of observed education outcomes rather than to average ones. For this reason, we use a Markov chain approach in which the transition matrix contains the overall transitions from high school grades to university grades at graduation. We also separate students coming from public and private high schools. This is done because there are significant differences between these types of schools related to tuition
configuration, grade inflation, each school’s student admittance criteria, etc.

In our opinion, an analysis of this kind has threefold relevance. First, the above-mentioned grades are relevant because they reflect human capital acquisition when young adults are about to enter the labor market (Betts & Morell, 1999; Krueger, 1999). As Miller (1998) shows, grades are positively related with productivity and, therefore, with earnings. Second, our approach allows us to compute the expected change on ergodic solutions derived from changes in transitions as a consequence of university policies. Finally, different university admission policies might be implemented for private and public high schools since different transitions are observed.

The remainder of this paper is structured as follows. The first section provides the theoretical application needed to compute the elasticity changes of the Markov chain transitions. The second section describes the data and shows the empirical evidence. The final section contains some concluding remarks.

Markov Chains: Computing Elasticities of Ergodic Solutions

In this section, we introduce the main characteristics of the Markov chains methodology applied to our study. First, we place high school grades in a finite state space, that is, we distribute students into several categories (states) according to their high school grades. At this stage, we wanted to establish the best criterion for discretizing Markov chain states and identify the ideal number of states of the chain. To aid the interpretation of the final results, we considered five states—low (L), middle-low (ML), intermediate (I), middle-high (MH), and high (H)—and defined them in such a way that each state contained a similar number of students. This step obliges us to define the limit values (boundaries) for each state. Therefore, we used six limit values to delimit these five categories: a lowest value, a highest value, and four intermediate values. The same limiting values for forming categories were set for the columns as for the rows.

Second, we compute the transition probabilities which constitute the elements of the transition matrix (M). Each transition probability indicates the probability that a student in the state \( i \) remains in \( i \) or transits to another state \( j \). Let’s suppose an individual is classified in the low-grade state (L). The transition matrix shows the probability that this student remains in this state \( (p_{iL}) \) as well as the probabilities to transit to the rest of the grade states \( (p_{iML}, p_{iLI}, p_{iLMH}, p_{iLH}) \). The same applies for the other transitions from the rest of the states. A homogenous Markov chain is characterized by probabilities of transiting from one state to another being independent of the time at which the step is made. Taking into account a model of grades in different moments in time, it means that the academic performance at the university \( (t) \) is dependent on the performance in high school \( (t-1) \) but independent of prior performances \( (t-2, t-3 \) and so on). This is the case when using first-order transitions (the immediately previous moment of time, \( t-1 \)). Therefore, Markov chains analysis allows us to calculate gains derived from an (educational policy) intervention at specific periods of time.

Third, results from the transition matrix allow us to compute the ergodic solution \( (\pi_{t+k}) \), that is, the probabilities of individuals to fall into each state of the chain in the long run. Note that the ergodic state is instructive for pedagogical purposes since students are in the matrix for only one time-cycle. They enter as freshmen (row) and leave after one time-period as graduates (column). This ergodic solution is calculated through expression (1). This implies raising the transition matrix to a large power \( (k) \), which is a considerable number of periods so as to obtain the long run solution.

\[
\pi_{t+k} = (M \times M \times M \times \ldots \times M) \pi_t = \pi_t M^k
\]  

(1)

The ergodic solutions are obtained from the eigenvector associated with the second eigenvalue of the transition matrix, which is equivalent to the vector obtained as \( k \) (iterations) increases and each row in the matrix converges on the ergodic solution. A concentration in a concrete state allows the inference that a long-run solution would show
convergence towards a common position in the distribution. For instance, when our final solution exhibits a probability of 0.8 at the first state of the chain, a considerable frequency of university students will be finally allocated into the lowest academic performance level.

Finally, inference becomes possible since Markov chains allow us to determine the changes in the ergodic solution when a change in transition probabilities occurs. According to Conlisk (1985), the effects of a change in transition probabilities in the ergodic solution of a chain \( (\pi) \) can be quantified. Let us suppose that there is a gain in the probability of a person remaining in the same category, that is, a state's persistence \( (\Delta p_{ii}) \). Clearly, this would bring about a loss in probability in, at least, one state in the chain where individuals could transit from \( i \) \( (p_{ij} \text{ being } i \neq j) \). Suppose that \( \varepsilon \) is the probability gain in one of the probabilities of the transition matrix \( (M) \). To compute the effect of the latter on the ergodic solution, the partial derivate \( \partial \pi / \partial \varepsilon \) must first be calculated, and the effect of readjustments in the transitional matrix on the ergodic solution can be estimated.

Only these derivatives allow the computation of real changes in the ergodic solution. The effects of the perturbation on the ergodic solution, as well as those on the mean of the first passage matrix, are not characterized in terms of discrete changes, but rather in terms of the direction of change represented by the derivatives, evaluated at \( \varepsilon = 0 \) (Conlisk, 1985). To compute this, the following sequence must be followed:

\[
Z = (I - M + u' b)^{-1} \quad \text{where } \pi = b' Z \quad (2)
\]

\[
\partial Z / \partial \varepsilon = Z E : Z \quad \text{where } \partial M / \partial \varepsilon = E \quad (3)
\]

\[
\partial \pi / \partial \varepsilon = b (\partial Z / \partial \varepsilon) \quad (4)
\]

in which \( u \) is a vector of ones with size \( 1 \times n \) (where \( n \) is the number of states), \( b \) is a vector that satisfies \( b' u = 0 \), \( I \) is the identity matrix, \( Z \) is what Conlisk (1985) defines as the “fundamental” matrix (there being a separate matrix for each choice of \( b \)), and \( E \) arises out of the changes that are made in the estimated transition matrix. In this application, \( b \) is selected as the initial probability vector which has been constructed assuming a discrete uniform distribution (with equiprobability for each state).

Ergodic solutions are sensitive to small changes in individual entries in the transition matrix. However, it is not immediately clear what should be considered small in such circumstances. It might be thought that since 0.1 or 0.01 are small absolute numbers, changes of such a degree would also be statistically insignificant. Yet, such a figure might be twice the previous value, and thus be highly significant for an individual’s achievement. Viewed in this light, the sensitivity analysis applied is, in fact, extremely informative.

### Data

The data for determining the column in the Markov transition state matrix correspond to the final grades obtained by students at the University of Barcelona for the period 1996–2003. Hence, our sample considers those students (named licenciados) who completed a university degree lasting five years (or more). The sample comprises 22,364 students from 14 of the largest faculties (representing 74.9% of all licenciados from the University of Barcelona). In terms of sample size, then, the study meets the requirements laid down for analyzing university grades (Horowitz & Spector, 2005; McNabb, Pal & Sloane, 2002; Smith & Naylor, 2005). In addition, from the degree courses considered, our sample represents 45.9% of all graduates in Catalonia (the Spanish Autonomous Community where the University of Barcelona is located and from which over 90% of its graduates come).

Regarding high school grades which were used to establish the row of the student, we use PAU grades (“PAU” stands for Prueba de aptitud para el acceso universitario), which is a weighted average of the national standardized examination which students take in order to enroll at a university (that accounts for 60% of the PAU grade) and their high school record (the latter comprises the average of the final grades of the last three courses of upper-secondary education and accounts for 40%). Related to university results, we use final grade,
which is not the result of a unique exam at the end of a university degree, but a weighted average of the grades of all subjects completed while at the university. It is weighted because exam results have a different value if students pass the subject in the first test or later (the value of the grade decreases at each exam attended).

Empirical Evidence

Table 1 shows the limit values to construct the row categories (states), the ergodic solutions, the transition matrix, and the changes in the ergodic solutions once a change in transition probabilities occurs. Results are differentiated based on the kind of high school attended (public, private-religious, or private-non-religious). Related to the limit values, they are constructed assuming the same initial frequency. Thus, values are different for each kind of school because each state contains 20% of the high school grades distribution. Note that grades range from 1 to 4.

The ergodic solutions display the long run probabilities for each state of the chain. These steady state results show dissimilarities, which are caused by the fact that we observed the dynamics of three different grade distributions which depend on the high school attended. Hence, these solutions are not strictly comparable in terms of grade values because each transition matrix considers its own distribution dynamics. Students having attended public high school show a greater long-run probability to achieve a higher academic performance than the lowest state (as the lower probability at the “L” state shows). Whereas, in the long run 78.72% of students from private religious schools and 76.43% of those in private non-religious schools will remain in the lowest state, only 60.58% of students in public schools would do so. The same can be stated through the examination of particular probabilities in the transition matrix. In short, the probability denoting those who remain in the lowest state of the chain (\( p_{LL} \) denoting no transition or transition from Low to Low) is lower for the transition matrix of those having attended a public school.

Finally, we determine the changes in the ergodic solution when a change in transition probabilities occurs. What might the effects be of an educational policy measure that increases the probability of students achieving higher university grades? Let us suppose that there is a 5% gain in the probability of a state’s persistence (\( \Delta p_{ii} = 0.05 \)) in the transition matrix. Clearly, if we consider that we follow up an improvement in one state, this would cause a loss in probability in the lower contiguous state. For example, a 5% gain in \( p_{ML} \) will suppose a 5% reduction in \( p_{MLML} \). We only take into account transitions from lower states to upper ones. Then, based on these changes in transition matrix, in Table 1 we show the changes between the initial ergodic solution and the one obtained once an increase has been assumed.

Results show significant changes in the ergodic solutions for the lower grades. Thus, a 5% increase in the transition from the lower state to the following level, as a consequence of an educational policy, generates between a 3–4% expected change in the ergodic solution (depending on the kind of school attended). Specifically, for those having attended a public high school, we observe a change of 3.00% from L to ML, whereas it is a 3.99% for those in private-religious centers and 3.77% for students in private non-religious schools. For lower stages, the highest effects of improvement in the ergodic solution are found for the students attending either religious or non-religious private schools (a difference of around one percentage point). Nonetheless, if we consider the upper states of the Markov chain, we observe higher effects for students from public schools although these are insignificant. Thus, the change in the ergodic solution for the upper state (H) is 0.33% for students from public schools whereas 0.19% and 0.20% are obtained for those in private-religious and private non-religious, respectively.

To sum up, educational policies that cause 5% increases in grades are more effective in the lower levels of the grades’ distribution as well as for students who come from private schools (religious and non-religious).
Table 1
Variations in Ergodic Solution According to High School Attended

<table>
<thead>
<tr>
<th>Public</th>
<th>Transition matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$L$</td>
</tr>
<tr>
<td>Ergodic solution:</td>
<td></td>
</tr>
<tr>
<td>[ \begin{pmatrix} 0.6058, &amp; 0.2328 &amp; 0.0953 &amp; 0.0519 &amp; 0.0143 \ L &amp; ML &amp; I &amp; MH &amp; H \end{pmatrix} ]</td>
<td>\begin{pmatrix} 0.6299 &amp; 0.2313 &amp; 0.0848 &amp; 0.0463 &amp; 0.0077 \ 0.6315 &amp; 0.2224 &amp; 0.0974 &amp; 0.0357 &amp; 0.0130 \ 0.5529 &amp; 0.2448 &amp; 0.1127 &amp; 0.0710 &amp; 0.0185 \ 0.4266 &amp; 0.2836 &amp; 0.1522 &amp; 0.1074 &amp; 0.0301 \ 0.1680 &amp; 0.1978 &amp; 0.1813 &amp; 0.2253 &amp; 0.2276 \end{pmatrix}</td>
</tr>
</tbody>
</table>

Limit values: (1, 1.5, 1.73, 1.95, 2.3, 4)

Changes in ergodic solution

ML: (-0.0303, \textbf{0.0300}, 0.0004, -0.0003, 0.0002)
I: (-0.0011, -0.0114, \textbf{0.0119}, 0.0005, 0.0001)
MH: (-0.0007, 0.0002, -0.0045, \textbf{0.0050}, 0.0001)
H: (-0.0011, -0.0002, 0.0002, -0.0021, \textbf{0.0033})

<table>
<thead>
<tr>
<th>Private religious</th>
<th>Transition matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ergodic solution:</td>
<td></td>
</tr>
<tr>
<td>(0.7872, 0.1158, 0.0592, 0.0305, 0.0073)</td>
<td>\begin{pmatrix} 0.8088 &amp; 0.1080 &amp; 0.0540 &amp; 0.0248 &amp; 0.0044 \ 0.7846 &amp; 0.1194 &amp; 0.0597 &amp; 0.0306 &amp; 0.0058 \ 0.7130 &amp; 0.1479 &amp; 0.0732 &amp; 0.0542 &amp; 0.0117 \ 0.5268 &amp; 0.2258 &amp; 0.1346 &amp; 0.0825 &amp; 0.0304 \ 0.1988 &amp; 0.1713 &amp; 0.1829 &amp; 0.2351 &amp; 0.2119 \end{pmatrix}</td>
</tr>
</tbody>
</table>

Limit values: (1, 1.58, 1.79, 2.01, 2.40, 4)

Changes in ergodic solution

ML: (-0.0405, \textbf{0.0399}, 0.0003, 0.0003, 0.0001)
I: (-0.0005, -0.0056, \textbf{0.0059}, 0.0002, 0.0000)
MH: (-0.0007, 0.0003, -0.0028, \textbf{0.0031}, 0.0001)
H: (-0.0008, 0.0000, 0.0001, -0.0012, \textbf{0.0019})

<table>
<thead>
<tr>
<th>Private non-religious</th>
<th>Transition matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ergodic solution:</td>
<td></td>
</tr>
<tr>
<td>(0.7643, 0.1506, 0.0408, 0.0329, 0.0114)</td>
<td>\begin{pmatrix} 0.7891 &amp; 0.1438 &amp; 0.0288 &amp; 0.0288 &amp; 0.0096 \ 0.7692 &amp; 0.1292 &amp; 0.0677 &amp; 0.0277 &amp; 0.0062 \ 0.6948 &amp; 0.1883 &amp; 0.0747 &amp; 0.0390 &amp; 0.0032 \ 0.4586 &amp; 0.3312 &amp; 0.1146 &amp; 0.0732 &amp; 0.0223 \ 0.1635 &amp; 0.2372 &amp; 0.1603 &amp; 0.2404 &amp; 0.1987 \end{pmatrix}</td>
</tr>
</tbody>
</table>

Limit values: (1, 1.55, 1.8, 1.98, 2.03, 4)

Changes in ergodic solution

ML: (-0.0390, \textbf{0.0377}, 0.0015, -0.0001, -0.0002)
I: (-0.0006, -0.0071, \textbf{0.0076}, 0.0001, 0.0000)
MH: (-0.0006, 0.0003, -0.0019, \textbf{0.0021}, 0.0000)
H: (-0.0008, 0.0000, 0.0002, -0.0013, \textbf{0.0020})

Note. The states refer to low academic achievement (L), middle-low (ML), intermediate (I), middle-high (MH), and high (H). As an example, changes in ergodic solution values for ML state represent the expected changes in the five probabilities of the ergodic solution assuming a 5% improvement in the transition matrix probability from the L state to the ML state.
Conclusion

Results show small changes in the ergodic solution if transition probability of grades is increased, at least for the considered period. Thus, as results show, a 5% increase in the transition matrix from the lowest state in high school grades to the immediately upper state (when considering university grades) raises no more than a 4% increase in the ergodic solution. In addition, for upper states, the gain in grades is much lower (less than 0.34%). Likewise, grades transition from high school to university differs according to the type of high school attended (public, private religious, or private non-religious). Specifically, transition from high school to university seems to be easier for students in public secondary schools.

It seems that the educational policies to be implemented should differ slightly in public and private schools, in view of the differences in transition performance observed in the students, and that they may specially improve transitions of students in lower states. This would have an incidence on human capital acquisition that may result in a better performance in the labor market.

Acknowledgements

The authors would like to thank the Servei de Gestió Acadèmica of the University of Barcelona for supplying the data and Ms. Amparo Asensi. Any errors, however, are attributable to the authors. Toni Mora gratefully acknowledges the financial support of the Spanish Ministry of Science and Technology by means of grant SEJ2006-01161/ECON.

Editor's Notes:

Those who read the preceding IR Apps, *Modeling Potential Implications of a Change in Tenure Policy: A System Dynamics Approach*, were reminded that one of the methodologies for looking at faculty flow is to use a Markov model to simulate the flow of faculty and the steady state of tenure. It can also be used to estimate the expected length of time a faculty member will be at a college when they start there in a specific category. It is truly a very fortunate event that this IR Applications is a demonstration of the use of a Markov model. It should be noted that this is a somewhat different use of Markov. In the use in faculty flow modeling there is a state that represents leaving the process. In this case of a Markov chain, there is no external state. While the Markov chain does not have the external state, it does have the property of having a steady state. This is a key property since it sets an absolute criterion. Given the matrix, in the long run, there is a steady state that is a unique characteristic of the chain.

The model developed was used to look at the implications of admitting students from different types of high schools: private or public. The methodology is also used to look at the change in an end-state given a change in the transition-state probabilities. This demonstrates one of the really important aspects of the Markov methodology—the ability to do *What-ifs?*

It should be noted that this use of the Markov creates a means for looking at various implications of changes in the enrollment of students. It does this by iterating entering and graduating categories, something that does not happen in the physical environment.

One of the more interesting and valuable aspects of this article comes from its threefold purpose. It is also in this purpose the reader is encouraged to think of extensions and next steps that can be taken because of the groundwork laid by the authors. First, there is the understanding that grades are related to value from the human capital perspective. One interesting next step would be to look at the careers and the earnings of those who graduate with different levels of grades. The second purpose is to look at shifts in the steady-state solutions with shifts in the transition matrix. This issue forms the basis for considering what university strategies or policies might cause such a shift. Considerations might include academic programs at the high school as
well as at the university. The third purpose is the revealed difference in transition matrices based on the type of high school. Is this a difference in grading patterns or does it represent a difference in student ability? It is possible that if the analysis were done separately for the grades and again for the standardized national test then the resulting differences in the transition matrix and the ergodic states might suggest the source of differences and the extent to which students from the public schools seem to do better given their entering scores. But then how might one construct a transition matrix with similar rows and columns?

Underlying all of these further analyses is the degree to which the students graduating from the university are a comparable sample from the three types of high schools. Do students from the types of high school have the same probability of attending the university? Do those who attend have the same probability of graduating from the university? How would the Markov analysis be extended to have external states such as not attending the university, leaving without graduating from the university, and graduating?

This article, as with all valuable research, raises questions as well as answers them. One final note for those who use the MMULT function in Excel, the transition-state matrices converge on the ergodic state by the 16 power.

References


IR Applications is an AIR refereed publication that publishes articles focused on the application of advanced and specialized methodologies. The articles address applying qualitative and quantitative techniques to the processes used to support higher education management.

Authors can submit contributions from various sources such as a Forum presentation or an individual article. The articles should be 10-15 double-spaced pages, and include an abstract and references. Reviewers will rate the quality of an article as well as indicate the appropriateness for the alternatives. For articles accepted for IR Applications, the author and reviewers may be asked for comments and considerations on the application of the methodologies the articles discuss.

Articles accepted for IR Applications will be published on the AIR Web site and will be available for download by AIR members as a PDF document. Because of the characteristics of Web-publishing, articles will be published upon availability providing members timely access to the material.

Please send manuscripts and/or inquiries regarding IR Applications to Dr. Gerald McLaughlin.