

A Compact Model with Longitudinal Data for the Analysis of School Transitions

Un Modello Compatto con Dati Longitudinali per l'Analisi dei Flussi Scolastici

Francesco Maccari

Università degli Studi di Firenze, Viale Morgagni 59, Firenze, maccari@ds.unifi.it

Giulio Ghellini

Università degli Studi di Siena, Piazza S. Francesco 8, Siena

Riassunto. La fase di riforma che sta investendo il sistema scolastico del nostro Paese in questi ultimi anni impone al ricercatore sociale nuovi e innovativi approcci allo studio delle carriere scolastiche, anche al fine di permettere un corretto monitoraggio dell'impatto che tali riforme hanno sul processo di formazione e di selezione degli studenti. In questo lavoro, utilizzando un'originale base di microdati longitudinali sulle carriere degli studenti della provincia di Pisa, si propone un modello di analisi per coorti in grado di tenere conto sia di variabili di contesto, sia della storia scolastica pregressa. Sulla base del modello proposto sono derivati alcuni indicatori sintetici sulle carriere e sugli esiti, la cui analisi evidenzia la persistenza di rilevanti differenziazioni socio culturali nei percorsi sperimentati.

Keywords: cohort study, transition models, education, student career indicators.

1. Introduction

The ultimate purpose of the paper is the presentation of a compact model for the analysis, in the framework of cohort studies, of student careers in the secondary school system. In general, the model is suitable to describe the whole educational system's dynamics, in terms of transition and exit matrices, but the data available at the moment refer only to high secondary school (*scuola secondaria superiore*).

Our approach recalls studies representing the educational system by means of a discrete finite Markov chain model (Bartholomew, 1982, Trivellato, 1980, Bernardi e Trivellato, 1980), but, instead of using contemporary stock data, builds up transition matrices for cohorts of students defined on the basis of the year of enrolment for the first time in the considered school system. This choice seems useful both for the opportunity to introduce explanatory variables in the analysis (gender, previous school career, family background, etc.) and to relax some constraints of transition models based on contemporary data.

Following the dynamic structure of the cohort (see Legaré, 1972), the concept of school path has been introduced and formalized. The model presented goes beyond Legaré's proposal, since it introduces the previous school path in the definition of all matrices, and, consequently, in the definition of a set of indicators aimed to monitor the system outcome. The work ends with some analyses of a set of performance indicators, showed

not only in their temporal evolution, but also comparing the behaviour of subgroups of students.

2. Data and model specification

The analysis we will show in this paper involves the high secondary school system of Pisa province. The data we have used refer solely to students enrolled in public high schools from school year 1993/94 to 2002/03. That makes a total of ten years of observations. More specifically, we used the following two data-sets:

- Entry variables: gender, age of enrolment and parents education level;
- School career: the record of every school year spent by each student into the system considered (grade of enrolment and final result).

The objective of the analysis, briefly recalled in the introduction, is to define indicators of students' performance, under different aspects, within a cohort approach.

The model we use to obtain indicators has been proposed only this year (see Maccari 2006 for a thorough treatment) and is based on the concept of path: a school path is defined as the set of all the possible ways in which a student with r_b past repetitions can progress from an initial school grade g_b to a final one g_e in y school years. Given this definition, a path is exactly defined once the set of parameters (g_b, r_b, g_e, y) has been fixed.

The model here recalled is able to measure the observed probability of all possible school paths in the cohort, using a recursive approach arising from the fact that every path can be decomposed into several nested sub-paths, and, ultimately, into several possible transition sequences that satisfy the fixed parameters. Hence, the matrix **W** of school paths probabilities is obtained by means of a correct walkthrough of matrix **P** of transition probabilities for the considered cohort.

Once matrix **W** is specified, a set of cohort performance indicators and tables is proposed, referred to the following aspects of school careers:

- Staying at school indicators, consisting of regularity tables and permanence tables;
- Indicators of attainment of different education levels: tables of regular exit with different titles and tables of exit in any number of years with different titles;
- School life expectancy indicators, which comprehend the expected regular school life and the expected permanence at school.

We want to underline here that all the indicators proposed are obtained as a composition of school paths probabilities and, hence, through the values of **W** matrix; as for the second class of indicators above mentioned, these values have to be composed with the correct probabilities of exit from school (Maccari, 2006, for the complete formalization).

3. Educational indicators analysis

All the analyses concentrate on indicators obtained by the already presented model. In particular, we take into account the following:

- Regularity table: is referred to a cohort starting at time T (usually fixed at 10.000 units) and reports how many units of the cohort have had a regular school career

- after k years;
- RED = Regular Exit with Diploma, drawn from the corresponding table of regular exit;
- ERSL = Expected Regular School Life;
- ED = Exit rate with Diploma in any number of years, obtained from the corresponding table of exit;
- EWT = Exit rate Without any Title in any number of years, obtained from the same table as ED;
- EPS = Expected Permanence at School.

The indicators have been calculated on two specific cohorts: the first one is 1994/95 cohort (N = 3226), since it gave more guarantees of data quality, with respect to 1993/94, and has been observed for a long enough timespan to follow the totality of experimented paths; the second one is 1998/99 cohort (N = 2856), which is the last providing for a minimum period of five years of observations, which corresponds to a regular career, and is also separated enough in time from the other one to look for significant changes in students' performance.

Looking at Table 1, it is possible to underline two different aspects. On one side, a better performance of the 1998/99 cohort arises from the following results: higher figures in regularity table, an increase in regular exit with diploma (RED) from 57,8% to 62,8% and a higher regular school life expectation (ERSL). On the other side, school performance is dramatically worse for students repeating the first year in secondary school (see cohort 1994/95): the exit rate with diploma (ED) declines from 70.7% to 24.0%, while the exit rate without any secondary school title (EWT) grows from 25.1% to 71.7%. Also the values of ERSL and RED follow the same pattern, confirming the utmost importance held by the result of first year in giving rise to students' school career in secondary education.

Table 1. Indicators of performance - Cohorts 1994/95 and 1998/99.

	Total students enrolled (T)		Students repeating first grade (T+1)	
	94/95	98/99	94/95	98/99
Regularity table				
Initial cohort	10.000	10.000	10.000	10.000
After 1 year	8.339	8.606	5.522	5.279
After 2 year	7.300	7.687	3.699	3.364
After 3 year	6.545	7.019	2.609	2.364
After 4 year	6.092	6.457	1.877	1.629
ERSL (years)	3,83	3,97	2,37	-
RED (%)	57,8	62,8	16,9	-
ED (%)	70,7	-	24,0	-
EWT (%)	25,1	-	71,7	-

Considering only cohort 94/95 (see Table 2), we want to point out some very interesting differences in school pattern among sub-groups of students with different values of entry variables (gender, age of enrolment, parents' education level) in the analysed system.

As for gender, females achieve a better performance than males as concerns regularity; indeed, they report a much higher value both in ERSL and in RED. We also want to underline the very high rate of graduates (ED close to 80%) and the marginal presence of female students dropping out without any title (EWT = 17.3%).

The age of entry in secondary education seems to be the variable that most influences

future student performance: all the indicators reported show that students entering the considered school system in “regular” age (≤ 14 years) achieve much better results than the ones entering later. In particular, the two exit rates with diploma (RED and ED) are well more than double, with respect to the other group. We also notice that the major part of students entering the school in non-regular age drop out with no title (EWT = 59.9%).

The last characteristic considered, the parents education level, lights up the differences still existing in the school careers of students due to a different family background. In our analyses, we considered the higher education level within the two parents, that is subdivided into two groups: high education level (that is high school diploma or university degree) and low education level (less than a high school diploma).

All the indicators presented show a better performance for students belonging to the first group, as can be seen from the table. The last column refers to students whose parents’ education level information is not available (about 12.1% of the cohort).

Table 2. *Indicators of performance by students characteristics – Cohort 1994/95*

	Gender		Age of entry		Parents education level		
	M	F	Regular	Later	High	Low	Missing
ERSL (years)	3,51	4,13	4,04	2,57	4,22	3,62	3,46
ESP (years)	4,69	4,86	4,95	3,73	4,93	4,72	4,49
RED	47,3	68,1	63,4	24,1	70,6	50,6	48,6
ED	61,7	79,5	77,2	31,8	81,3	64,8	62,4
EWT	33,1	17,3	19,3	59,5	16,5	29,3	33,7

4. Conclusions

The model proves to be suitable to catch cohort dynamics and to exploit the added value given by longitudinal data. Indicators are useful to show changes in performance in time and different behaviours among groups of students. We think that a desirable evolution would be taking into account, in the building of matrices, also the transitions of students among different school types, which is an important component of the overall dynamics.

References

- Bartholomew D. J. (1973) *Stochastic Models for Social Processes*, Wiley, Chichester.
- Bernardi L., Trivellato U. (1980) Un modello markoviano del processo scolastico: (II) suo impiego per l’analisi della selettività del sistema preuniversitario italiano, *Rivista di Statistica Applicata*, 2, 55-90.
- Legaré J. (1972) Methods for measuring school performance through cohort analysis, *Demography*, 4, 617-624.
- Maccari F. (2006) *Following Students: a Compact Cohort-Based Model with Longitudinal Data for Student Careers Careers and Educational System Analyses*, PhD Thesis, Dipartimento di Statistica “G. Parenti”, Firenze.
- Trivellato U. (1980) Un modello markoviano del processo scolastico: (I) specificazione e procedimento di stima, *Rivista di Statistica Applicata*, 1, 3-20.