Sequence analysis

Hugo Harari-Kermadec

with Leonard Moulin

November 7, 2019

M2 EPOG - University Paris 13 - Econometrics

TraMineR

- TraMineR is an R package for sequence analysis.
- Specially designed for the social sciences.
- TraMineR: Trajectory Miner in R (Gewurztraminer wine ?)
- Freely available on the CRAN (Comprehensive R Archive Network)
  http://cran.r-project.org/web/packages/TraMineR/
- `install.packages("TraMineR")`
- `library("TraMineR")`

Sequence analysis

- In the social sciences: survival analysis $\Rightarrow$ using different types of modeling (mainly risk models and survival models) we focus on a particular event (marriage, first job, unemployment...).
- Development approach in terms of life course: whole trajectory (possibly with multiple transitions) as the unit of analysis

$\Rightarrow$ Holistic approach.

Sequence analysis

- In social sciences, sequences represent trajectories:
  - professional;
  - school;
  - sentimental;
  - residential;
  - union;
- With sequence analysis, we can study recurring patterns in the trajectories, taking into account the multiplicity of feasible states.
- A state is a situation in which an individual is at any given time (single, cohabiting, married, divorced, widowed).
- Sequences of states: a pathway is a sequence of states ordered along a time axis (single → married → widowed).

Sequence analysis

- Which characteristics of sequences are we interested in?
- What kind of indicators can we compute for a sequence set?
- What are suited plots for rendering sequences?
- How can we measure similarity between sequences?

Sequence analysis

With a more analytical or explanatory concern, we also consider issues such as:

- How can we identify groups with similar patterns and build typologies of sequences?
- How can we analyze the relationship of sequences with covariates?

Sequence analysis

Questions in social sciences:

- Do life courses obey some social norm? Which are the standard trajectories? What kind of departures do we observe from these standards ? How do life course patterns evolve over time ?
- Why are some people more at risk to follow a chaotic trajectory or to stay stuck in a state? How does the trajectory complexity evolve across birth cohorts?
- How is the life trajectory related to sex, social origin and other cultural factors?

Illustrative data set

- Study by McVicar and Anyadike-Danes (2002) on the transition from school to work in Northern Ireland.
- Dataset included in TraMineR.
- The aim of the study:
  - Describe the typical transition of young Irish.
  - Identified problematical trajectories. Distinguish between successful and unsuccessful transitions.
  - Understanding the factors that influences the trajectories.
  - Identified groups of young irish who have more problems to enter into the labor market.

Illustrative data set

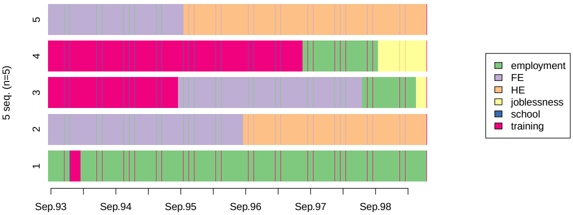
- 712 individuals
- Follow-up starting at the end of the compulsory education (July 1993)
- Time series of 70 status variables: September 1993 to June 1999.
- The alphabet is made of the following statuses: EM (Employment), FE (Further Education), HE (Higher Education), JL (Joblessness), SC (School), TR (Training).

State sequences

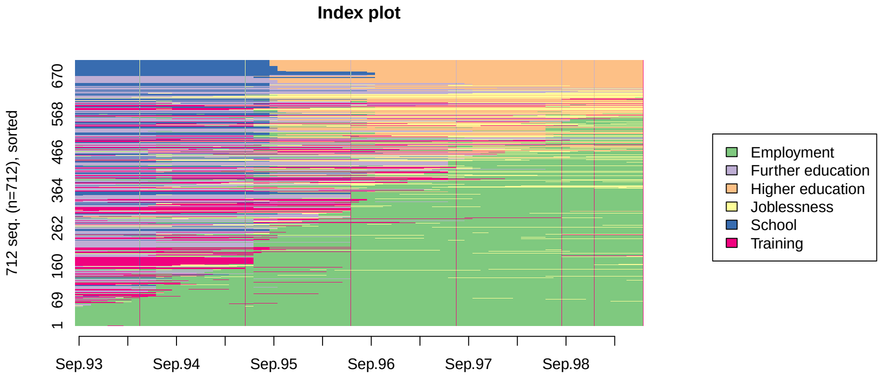
We define a state sequence of length l as an ordered list of l elements successively chosen from a finite set A. We represent a sequence x by listing the successive elements that form the sequence $x = (x_1, x_2, ..., x_l)$, with $x_j \in A$.

Plotting individual sequences: `seqiplot`

```
    Sequence
[5] (FE,25)-(HE,45)
[4] (TR,47)-(EM,14)-(JL,9)
[3] (TR,24)-(FE,34)-(EM,10)-(JL,2)
[2] (FE,36)-(HE,34)
[1] (EM,4)-(TR,2)-(EM,64)
```

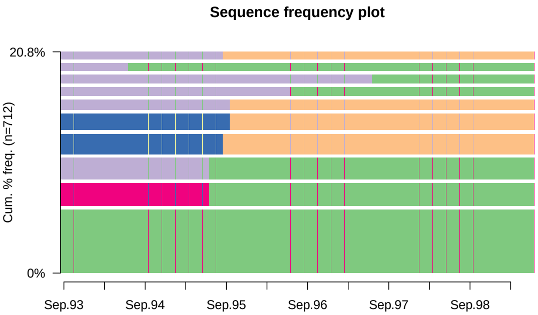


Carpet: seqiplot for the whole dataset

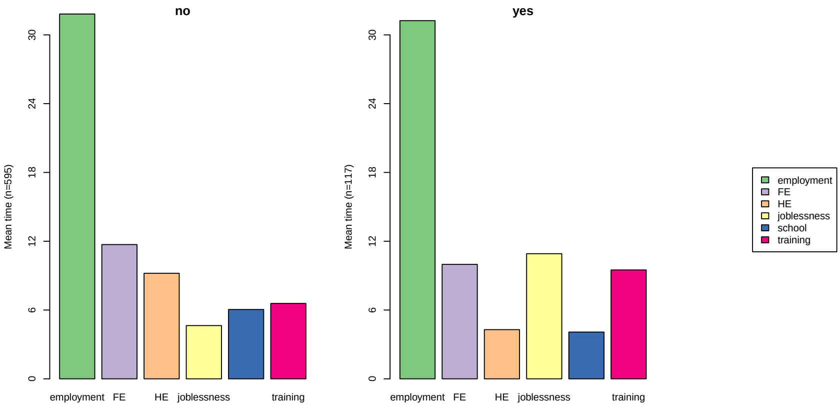


Sequences frequencies: seqfplot

	Freq	Percent
...		
FE/22-EM/48	14	2.5
TR/22-EM/48	18	2.5
EM/70	15	4.1



Mean time and covariables (unemployment): `seqmtplot`



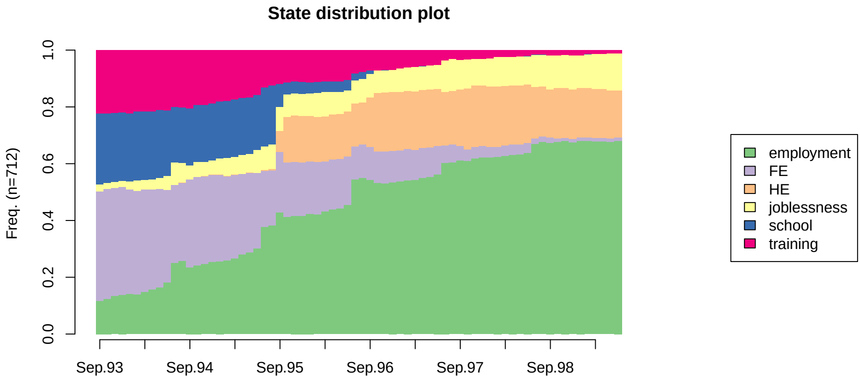
Transition rates: seqtrate

The transition rate between (s_i, s_j) is the probability to switch from state s_i to state s_j .

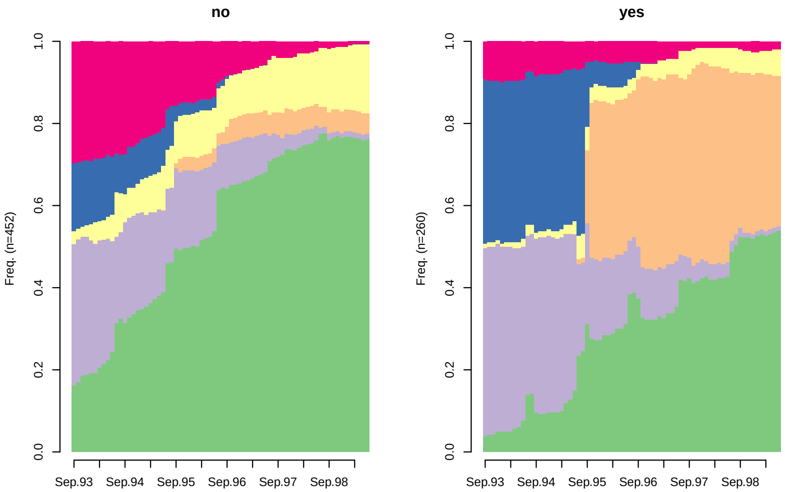
$$p(s_j|s_i) = \frac{\sum_{t=1}^{L-1} n_{t,t+1}(s_i, s_j)}{\sum_{t=1}^{L-1} n_t(s_i)}. \tag{1}$$

	[-> employment]	[-> FE]	[-> HE]	[-> joblessness]	[-> school]	[-> training]
[employment ->]	0.99	0.00	0.00	0.01	0.00	0.00
[FE ->]	0.03	0.95	0.01	0.01	0.00	0.00
[HE ->]	0.01	0.00	0.99	0.00	0.00	0.00
[joblessness ->]	0.04	0.01	0.00	0.94	0.00	0.01
[school ->]	0.01	0.01	0.02	0.01	0.95	0.00
[training ->]	0.04	0.00	0.00	0.01	0.00	0.94

Overall distribution: `seqdplot`



Distribution vs covariable: seqdplot(, group=honours)



1. Sequence analysis

2. Sequences typology
   - Measuring sequences (dis)similarities
   - Application

Measuring sequences (dis)similarities

- Methods from computer sciences (Hamming, 1950; Levenshtein,1966) and molecular biology. They appeared in the social sciences under the guidance of Abbott's seminal works.
- Methods based on the use of a measure of distance between sequences in order to observe similarities (and dissimilarities) between trajectories and build typology of sequences.
- Typology is used to identify and study the existing patterns in students pathways.

Measuring sequences (dis)similarities

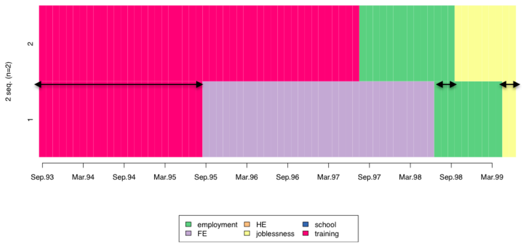
What is distance ?

- Most of advanced sequence analysis methods rely on a dissimilarity measure: how far two objects are.
- For instance, consider two incomes x and y:
  - $d(x, y) = |x - y|$.
  - $d(x, y) = log(1 + |x - y|)$.
  - $d(x, y) = (x - y)^2$.
- How to do it with categorical sequences?
- Depending on the issue, we want our dissimilarity measure to account for:
  - order of the states and transitions in each sequence;
  - temporality of the transitions;
  - duration of stay in each state;

- Choice of the measure is a crucial step $\Rightarrow$ depends i) data and ii) research question.
- 2 groups of measures:
  1. measures based on common attributes between sequences, i.e. measures that do not allow to shift part of the sequence (HAM, LCP, LCS);
  2. editing measures, i.e. measures taking into for similar shifted patterns (OM).
- Example: without shift, $x = ABAB$ and $y = BABA$ are very distant, while they are quite similar if we shift y by just one position.

Hamming distance

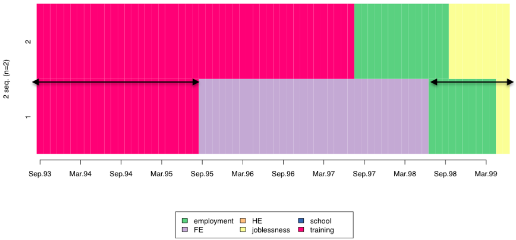
The simple Hamming distance (Hamming, 1950) is the number of positions at which two sequences of equal length differ.



Hamming distance focus on simultaneity: same state at same time.

LCS distance

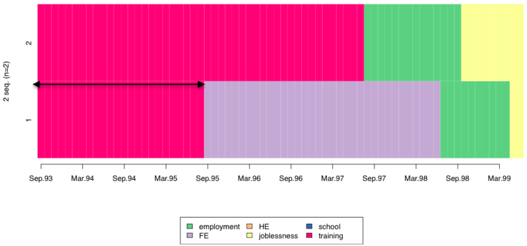
Total length minus 2× the length of the longest common subsequence (LCS), without time constraint



LCP distance

Total length minus 2× the length of the longest common prefix (LCP)

Figure: LCP - example on sequences 3 and 4 of mvad



Edit distance

An edit distance is defined as the minimal cost of transforming one sequence into the other. Two types of operations:

1. substitution of one element by an other
2. the insertion or deletion (indel) of an element, which generates a one position shift of all the elements on its right

Operations

Indel deforms the time structure to allow common subsequences to emerge.

Substitution conserves the time structure: elements are compared at constant date.

Edit distances

Example: indel

- Consider the two following sequences:

1	SC	SC	SC	EM	EM	EM	JL
2	SC	SC	SC	EM	EM	JL	JL

- Insertion of "EM", cost = 1:

1	SC	SC	SC	EM	EM	EM	JL	
2	SC	SC	SC	EM	EM	EM	JL	JL

- Insertion of "JL", cost = 1:

1	SC	SC	SC	EM	EM	EM	JL
2	SC	SC	SC	EM	EM	EM	JL

$\Rightarrow$ Two sequences are now identical, total cost = 2.

Edit distances

Example: substitution

- Consider the two following sequences:

1	SC	SC	SC	EM	EM	EM	JL
2	SC	SC	SC	EM	EM	JL	JL

- Substitution of "EM" by "JL", cost = 2:

1	SC	SC	SC	EM	EM	EM	JL
2	SC	SC	SC	EM	EM	EM	JL

⇒ Two sequences are now identical, total cost = 2.

Edit distances: Optimal matching

Let Σ the alphabet and λ the zero element. The different operations are:

- $a \to b$ an operation of substitution, with $a, b \in \Sigma \cup \{\lambda\}$ and $a \neq b$;
- $a \to \lambda$ an operation of delete; suppression;
- $\lambda \to a$ an operation of insertion.

$T_{x,y} = T_1 ... T_l$ denotes the set of l operations needed to transform a sequence x into a sequence y. We note $\gamma(T_i)$ the cost associated with the operation T_i.
We obtain the distance $d_{OM}(x, y)$ by calculating the sequence of operations $T_{x,y}$ which minimizes the total cost.

$$d_{OM}(x, y) = \min \left\{ \sum_{i=1}^{l} \gamma(T_i) \right\}. \tag{2}$$

Edit distances

Example: OM

- Consider the two following sequences:

1	EM	EM	FE	FE	HE	HE
2	FE	FE	HE	HE	EM	EM

- With a substitution cost = 2 and a indel cost = 1, the OM distance = 4 (on a maximum of 12):

1	EM	EM	FE	FE	HE	HE	-	-
2	-	-	FE	FE	HE	HE	EM	EM
Coût	1	2	2	2	2	2	3	4

Edit distances

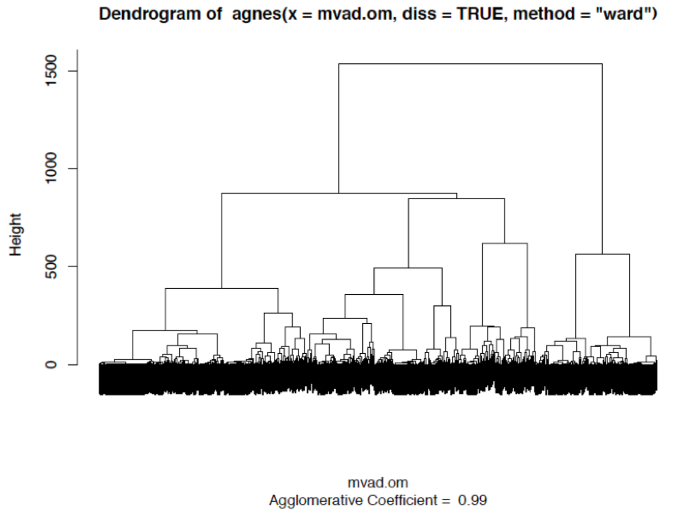
Costs

- Subsitution costs define the distance between two states *at the same time*.
- Indel costs defines the possibility of allowing a time lag in the comparison of sequences.
- If a substition cost is greater than twice indel cost, it will never be used.
- Ways to define subsitution costs:
  - **Theoretical costs:** costs are defined arbitrarily by the researcher. Some states may be closer to a state than others. Justification?
  - **Constant costs:** costs of substitution between states are constant.
  - **Estimated costs from the data:** costs estimated on transition rates between states.

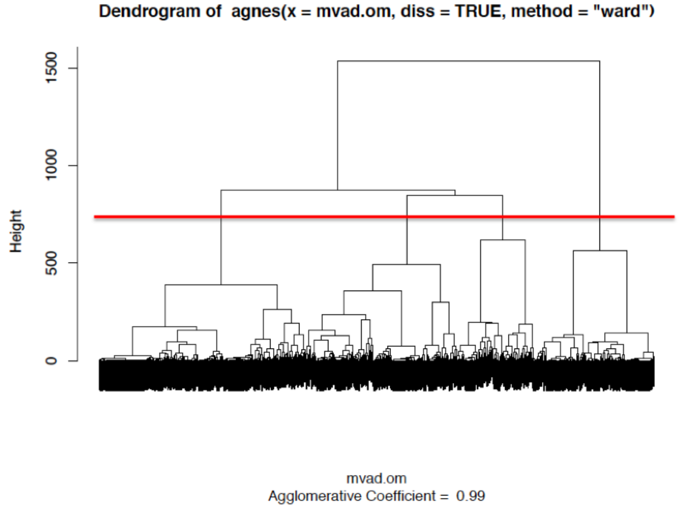
Classification

- Classification methods are used to construct a typology of sequences, i.e regroup the population of the sample into groups based on common characteristics between sequences.
- This grouping procedure is based on a simplification of the data.
- Describe reality?
- Two main types of clustering procedures:
  1. hierarchical clustering (ascending and descending);
  2. non hierarchical clustering (partitioning).

Hierarchical clustering: dendrogram



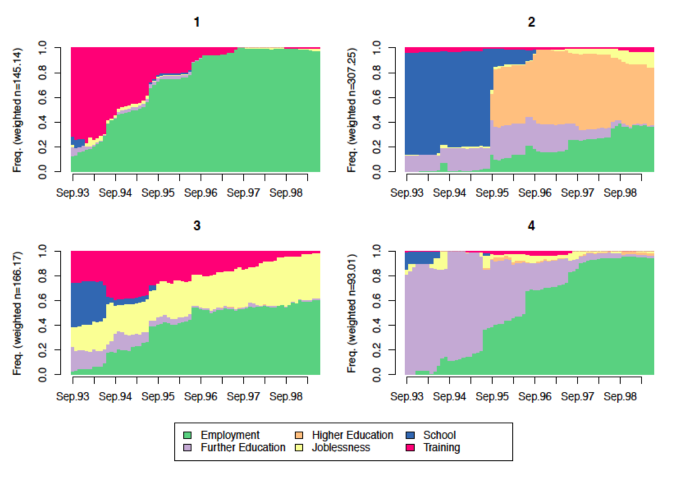
Hierarchical clustering: dendrogram



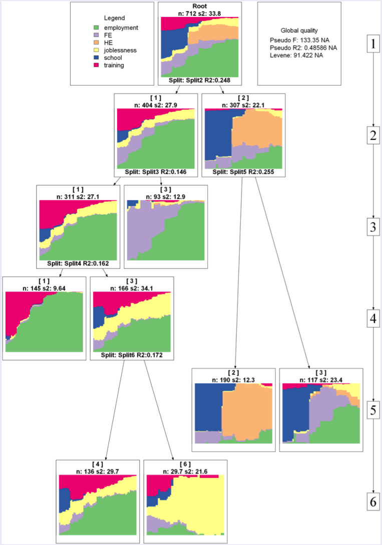
Clustering

Hierarchical clustering

Figure: Typical trajectories - ward



Hierarchical clustering: Ward Tree



Introduction

- The increase of tuition fees is a “political” solution adopted by many countries.
- France: universities autonomy (LRU) and statutes derogatory in some institutions (Dauphine, Science Po, Mines Telecom...).
- Analysis of the effects of the introduction of tuition fees at the University Paris Dauphine 9 on:
  1. students pathways selection ;
  2. student achievement.
- Methodology:
  1. Optimal matching to construct a typology of students trajectories in HE.
  2. Multinomial logit to assess the effect of tuition fees on the academic pathways selected by this university.
  3. Difference-in-differences in a non-linear model Puhani (2012).

Literature review

- The relationship between tuition fees and students' decisions has been extensively studied in the literature, in terms of:
    1. access to higher education (Cameron and Heckman, 2001; Coelli, 2009; Hubner, 2012...);
    2. choice of curriculum (Callender and Jackson, 2008; Dietrich and Gerner, 2012; Field, 2009...);

⇒ Persistent controversy of on the impact of tuition fees.

- Small number of empirical works on student achievement:
    1. effects of tuition fees on the time necessary to graduation (Garibaldi, Giavazzi, Ichino and Rettore, 2012);
    2. effects of tuition fees on student achievement (as endogenous variable, see Heineck, Kifmann and Lorenz, 2006).

⇒ The various impacts of the introduction of tuition fees put forward a debate on the existence, nature and extent of segregation phenomena.

Data

SISE

- The SISE database collects data on students.
- Data concerning enrollment and success of students in French universities;
- Data (more than 70 variables) about:
  - Sociodemographic characteristics of students: sex, age, social category, nationality, geographic origin, etc.
  - Schooling: establishment, diploma course, school attended, registration type, etc.
  - Previous schooling: bachelor academy, graduation year, baccalaureat, year of first registration in the French university system, success Diploma etc.
- Matching: SISE *universites - inscriptions*, SISE *universites - resultats* and ALGAE.

Data

Dauphine

- 1st university to implement tuition fees in France .
- This allows to analyze the effect induced by the increase of tuition fees on the students trajectories.

Gross income (per year)	Tuition fees
$<$ 40 000	1 500
40 000 - 50 000	2 000
50 000 - 60 000	2 500
60 000 - 70 000	3 000
70 000 - 80 000	3 500
$>$ 80 000	4 000

- Two cohorts:
  1. 2009-2010: all second year master in economy and management are without tuition fees;
  2. 2010-2011: a part of them increase tuition fees.

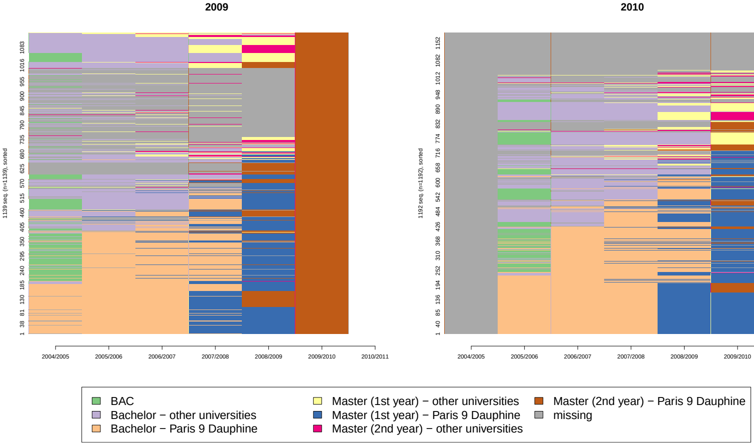
Table: Socio-economic characteristics of students in Master 2 economics-management at Dauphine by year and by type of Masters

Caracteristiques socio-economiques	2009 (n=1139) DN	2009 (n=1139) DGE	2009 (n=1139) Total	2010 (n=1192) DN	2010 (n=1192) DGE	2010 (n=1192) Total	Total (n=2331)
Social class							
Very well-off	56.03	60.17	58.21	59.34	60.48	59.98	59.12
Well-off	10.02	10.00	10.01	10.40	11.14	10.82	10.42
Average	14.84	14.33	14.57	17.34	16.64	16.95	15.79
Disadvantaged	19.11	15.50	17.21	12.91	11.74	12.25	14.67
Scholarships							
No	90.54	84.67	87.45	86.51	82.47	84.23	85.80
Yes	9.46	15.33	12.55	13.49	17.53	15.77	14.20
Sex							
Female	53.25	56.33	54.87	52.99	53.64	53.36	54.10
Male	46.75	43.67	45.13	47.01	46.36	46.64	45.90
Nationality							
French	69.20	79.33	74.54	68.40	81.43	75.76	75.16
Foreign	30.80	20.67	25.46	31.60	18.57	24.24	24.84
Location							
Paris	40.45	39.83	40.12	40.66	43.54	42.28	41.23
Seine et Marne	2.97	2.17	2.55	1.93	1.78	1.85	2.19
Yvelines	3.90	8.00	6.06	4.82	6.54	5.79	5.92
Essonne	2.04	3.83	2.99	4.05	3.27	3.61	3.30
Hauts de Seine	14.66	14.17	14.40	16.38	13.97	15.02	14.71
Seine Saint Denis	4.27	2.33	3.25	4.24	2.53	3.27	3.26
Val de Marne	7.42	6.50	6.94	7.13	5.35	6.12	6.52
Val d'Oise	2.41	2.00	2.19	3.85	3.86	3.86	3.05
Outside Ile-de-France	21.89	21.17	21.51	16.96	19.17	18.20	19.82
Age							
≤ 22	24.12	32.17	28.36	24.66	25.41	25.08	26.68
[23; 24]	48.24	46.33	47.23	46.24	57.80	50.06	50.06
≥ 25	27.64	21.50	24.41	29.09	16.79	22.15	23.25

Socio-economic characteristics	2009 (n=23 807)			2010 (n=25 910)			Total (n=49 717)
	Univ.	Dauphine	Total	Univ.	Dauphine	Total	
Social class							
Very well-off	30,7	58,2	32,0	30,7	60,0	32,1	32,0
Well-off	12,3	10,0	12,1	11,7	10,8	11,7	11,9
Average	18,5	14,6	18,3	18,0	17,0	18,0	18,1
Disadvantaged	38,6	17,2	37,6	39,5	12,3	38,3	37,9
Scholarship							
No	78,4	87,5	78,8	77,4	84,2	77,7	78,2
Yes	21,6	12,5	21,2	22,6	15,8	22,3	21,8
Sex							
Female	53,1	54,9	53,2	53,4	53,4	53,4	53,3
Male	46,9	45,1	46,8	46,6	46,6	46,6	46,7
Nationality							
French	61,8	74,5	62,4	59,6	75,8	60,4	61,3
Foreign	38,2	25,5	37,6	40,4	24,2	39,6	38,7
Location							
Paris	6,3	40,1	7,9	5,7	42,3	7,3	7,6
Île de France	13,2	38,4	14,4	12,4	39,5	13,7	14,0
Province	62,5	16,1	60,3	63,2	14,4	61,0	60,6
Out of France	18,0	5,4	17,4	18,7	3,8	18,0	17,7
Baccalauréat							
Littéraire	4,9	3,4	4,8	4,9	3,4	4,9	4,8
Économique	28,3	26,9	28,2	28,0	27,6	28,0	28,1
Scientifique	22,8	50,8	24,2	21,5	51,0	22,9	23,5
Technologique	6,7	1,1	6,4	6,5	0,9	6,3	6,3
Others technol.	1,8	0,1	1,7	1,6	0,2	1,5	1,6
Professionnel	0,7	0,3	0,6	0,5	0,1	0,5	0,6
Dispensé	34,9	17,6	34,1	36,9	16,8	36,0	35,1
Age							
≤ 22	22,5	28,4	22,8	22,9	25,1	23,0	22,9
[23; 24]	40,4	47,2	40,8	41,3	52,8	41,8	41,3
≥ 25	37,1	24,4	36,5	35,8	22,2	35,2	35,8

Methodology

Figure: Student trajectories depending on the year of enrollment in second year of master in economics and management at University Paris 9 Dauphine

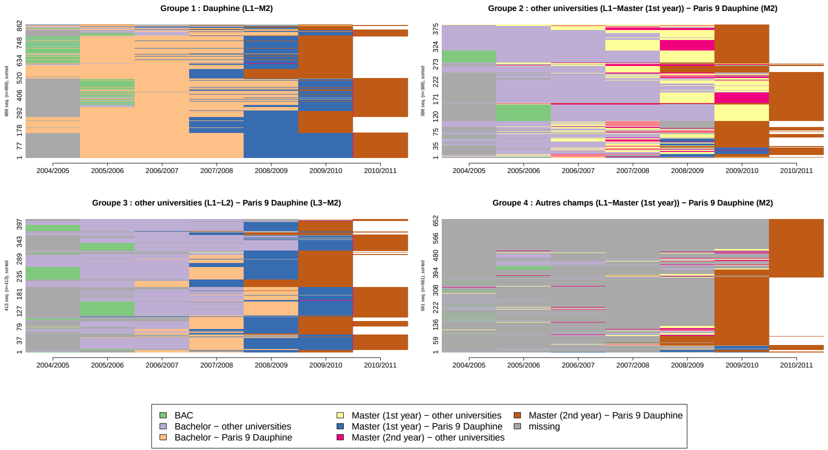


Optimal Matching Analysis

- OM with indel=max(costs)/2.
- Costs based on transition rates.

Results

Figure: Typical pathways of students in Master 2 at Dauphine



Results

Table: Situation in the previous year for the students in Group 4

Situation in previous year	2009 (n=346)			2010 (n=315)			Total
	DN	DGE	Total	DN	DGE	Total	(n=661)
University	38.55	17.22	27.46	43.06	22.81	32.06	29.65
Maneagement school	10.84	17.78	14.45	12.50	22.81	18.10	16.19
Engineering school	10.24	21.11	15.90	9.03	16.96	13.33	14.67
Foreign establishment	17.47	17.22	18.69	17.34	13.19	14.62	15.73
Other establishment in France	4.82	4.44	4.62	6.25	0	2.86	3.78
Other SISE establishment	0	0	0	0.69	0.58	0.63	0.30
Return to studies	18.07	22.22	20.23	15.28	22.22	19.05	19.67

Table: Socio-economic characteristics of students in the Master 2 economics-management at Dauphine by typical pathway

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
Social class					
Very well-off	67.55	50.26	56.90	54.61	59.12
Well-off	9.67	11.86	11.38	9.98	10.42
Average	13.46	18.81	16.46	16.64	15.79
Disadvantaged	9.32	19.07	15.25	18.76	14.67
Scholarship					
No	86.19	79.38	77.48	94.25	85.80
Yes	13.81	20.62	22.52	5.75	14.20
Sex					
Female	58.11	51.55	48.67	53.71	54.10
Male	41.89	48.45	51.33	46.29	45.90
Nationality					
French	87.11	83.51	80.87	50.98	75.16
Foreign	12.89	16.49	19.13	49.02	24.84
Location					
Paris	45.91	34.79	35.11	42.66	41.23
Seine et Marne	2.88	3.87	1.21	0.91	2.19
Yvelines	9.78	3.35	6.30	2.12	5.92
Essonne	3.57	2.84	4.12	2.72	3.30
Hauts de Seine	16.00	7.22	17.19	15.89	14.71
Seine Saint Denis	3.45	4.12	2.66	2.87	3.26
Val de Marne	5.75	8.25	6.54	6.51	6.52
Val d'Oise	3.68	4.38	2.66	1.66	3.05
Outside Ile-de-France	8.98	31.19	24.21	24.66	19.82
Age					
≤ 22	36.02	21.65	32.93	13.46	26.68
[23; 24]	53.97	52.32	55.21	40.39	50.06
≥ 25	10.01	26.03	11.86	46.14	23.25

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Very well-off (*ref.*)	-	-	-	-
Well-of	-0.064†	0.056†	0.016	-0.008
	(0.036)	(0.029)	(0.029)	(0.036)
Average	-0.090**	0.055*	0.000	0.035
	(0.031)	(0.024)	(0.025)	(0.030)
Disadvantaged	-0.155***	0.084***	0.015	0.056†
	(0.036)	(0.026)	(0.027)	(0.032)
Scholarship				
No (*ref.*)	-	-	-	-
Yes	0.023	0.052*	0.096***	-0.172***
	(0.034)	(0.025)	(0.024)	(0.038)
Sex				
Female (*ref.*)	-	-	-	-
Male	-0.057**	0.003	0.045**	0.009
	(0.022)	(0.018)	(0.017)	(0.021)
Nationality				
French (*ref.*)	-	-	-	-
Foreign	-0.192***	-0.082***	-0.000	0.275***
	(0.024)	(0.024)	(0.023)	(0.023)
Localisation				
Paris (*ref.*)	-	-	-	-
Ile de France	-0.025	-0.004	0.025	0.003
	(0.023)	(0.021)	(0.020)	(0.024)
Outside Ile de France	-0.349***	0.129***	0.053*	0.167***
	(0.033)	(0.023)	(0.024)	(0.028)
Age	-0.081***	0.025***	-0.027***	0.082***
	(0.008)	(0.005)	(0.006)	(0.006)
Type of Master's				
National (*ref.*)	-	-	-	-
Fee paying	0.039	-0.066***	-0.056*	0.083**
	(0.030)	(0.025)	(0.024)	(0.031)
Year				
2009 (*ref.*)	-	-	-	-
2010	0.002	0.0247	0.048*	-0.074**
	(0.028)	(0.022)	(0.022)	(0.028)

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Very well-off (*ref.*)	-	-	-	-
Well-of	-0.064†	0.056†	0.016	-0.008
	(0.036)	(0.029)	(0.029)	(0.036)
Average	-0.090**	0.055*	0.000	0.035
	(0.031)	(0.024)	(0.025)	(0.030)
Disadvantaged	-0.155***	0.084***	0.015	0.056†
	(0.036)	(0.026)	(0.027)	(0.032)
Scholarship				
No (*ref.*)	-	-	-	-
Yes	0.023	0.052*	0.096***	-0.172***
	(0.034)	(0.025)	(0.024)	(0.038)
Sex				
Female (*ref.*)	-	-	-	-
Male	-0.057**	0.003	0.045**	0.009
	(0.022)	(0.018)	(0.017)	(0.021)
Nationality				
French (*ref.*)	-	-	-	-
Foreign	-0.192***	-0.082***	-0.000	0.275***
	(0.024)	(0.024)	(0.023)	(0.023)
Localisation				
Paris (*ref.*)	-	-	-	-
Ile de France	-0.025	-0.004	0.025	0.003
	(0.023)	(0.021)	(0.020)	(0.024)
Outside Ile de France	-0.349***	0.129***	0.053*	0.167***
	(0.033)	(0.023)	(0.024)	(0.028)
Age	-0.081***	0.025***	-0.027***	0.082***
	(0.008)	(0.005)	(0.006)	(0.006)
Type of Master's				
National (*ref.*)	-	-	-	-
Fee paying	0.039	-0.066***	-0.056*	0.083**
	(0.030)	(0.025)	(0.024)	(0.031)
Year				
2009 (*ref.*)	-	-	-	-
2010	0.002	0.0247	0.048*	-0.074**
	(0.028)	(0.022)	(0.022)	(0.028)

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Very well-off (*ref.*)	-	-	-	-
Well-of	-0.064†	0.056†	0.016	-0.008
	(0.036)	(0.029)	(0.029)	(0.036)
Average	-0.090**	0.055*	0.000	0.035
	(0.031)	(0.024)	(0.025)	(0.030)
Disadvantaged	-0.155***	0.084***	0.015	0.056†
	(0.036)	(0.026)	(0.027)	(0.032)
Scholarship				
No (*ref.*)	-	-	-	-
Yes	0.023	0.052*	0.096***	-0.172***
	(0.034)	(0.025)	(0.024)	(0.038)
Sex				
Female (*ref.*)	-	-	-	-
Male	-0.057**	0.003	0.045**	0.009
	(0.022)	(0.018)	(0.017)	(0.021)
Nationality				
French (*ref.*)	-	-	-	-
Foreign	-0.192***	-0.082***	-0.000	0.275***
	(0.024)	(0.024)	(0.023)	(0.023)
Localisation				
Paris (*ref.*)	-	-	-	-
Ile de France	-0.025	-0.004	0.025	0.003
	(0.023)	(0.021)	(0.020)	(0.024)
Outside Ile de France	-0.349***	0.129***	0.053*	0.167***
	(0.033)	(0.023)	(0.024)	(0.028)
Age	-0.081***	0.025***	-0.027***	0.082***
	(0.008)	(0.005)	(0.006)	(0.006)
Type of Master's				
National (*ref.*)	-	-	-	-
Fee paying	0.039	-0.066***	-0.056*	0.083**
	(0.030)	(0.025)	(0.024)	(0.031)
Year				
2009 (*ref.*)	-	-	-	-
2010	0.002	0.0247	0.048*	-0.074**
	(0.028)	(0.022)	(0.022)	(0.028)

Conclusion

- Paris - Dauphine is the first university that introduced tuition fees in France.
- Education is a process, not a instantaneous decision.

$\Rightarrow$ Tuition fees have changed the types of students pathways allowing access to second year of master 2 and therefore the nature of people admitted in these curriculum $\Rightarrow$ Cumulative mechanism.