



HAL
open science

Characterizing Spanish Labour Pathways of young people with vocational lower-secondary education

Helena Corrales Herrero, Beatriz Rodríguez Prado

► **To cite this version:**

Helena Corrales Herrero, Beatriz Rodríguez Prado. Characterizing Spanish Labour Pathways of young people with vocational lower-secondary education. *Applied Economics*, 2011, pp.1. 10.1080/00036846.2011.581218 . hal-00712379

HAL Id: hal-00712379

<https://hal.science/hal-00712379>

Submitted on 27 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Characterizing Spanish Labour Pathways of young people with vocational lower-secondary education

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-2010-0174.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	01-Dec-2010
Complete List of Authors:	Corrales Herrero, Helena; Universidad de Valladolid, Economía Aplicada Rodríguez Prado, Beatriz; Universidad de Valladolid, Economía Aplicada
JEL Code:	J24 - Human Capital Skills Occupational Choice Labor Productivity < J2 - Time Allocation, Work Behavior, and Employment Determination/Creation < J - Labor and Demographic Economics, J64 - Unemployment: Models, Duration, Incidence, and Job Search < J6 - Mobility, Unemployment, and Vacancies < J - Labor and Demographic Economics, I21 - Analysis of Education < I2 - Education < I - Health, Education, and Welfare
Keywords:	school-to-work transition, sequence analysis, vocational training

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SCHOLARONE™
Manuscripts

For Peer Review

Characterizing Spanish Labour Pathways of Young People with Vocational Lower-Secondary Education

H. Corrales-Herrero^{a,*} and B. Rodríguez-Prado^a

^a*Departamento de Economía Aplicada, University of Valladolid*

Avenida del Valle de Esgueva, 6 47011 Valladolid, España

*Corresponding author. E-mail: helena@eace.uva.es

School-to-work transition is no longer considered to take place at one point but it is seen as a sequential process with multiple stages that extend over a relatively long period in which individuals gradually settle down into the labour market. This paper constructs and examines the labour pathways followed by Spanish young people who have completed vocational lower-secondary education. To do so, we use the Survey on Educational-Training Transition and Labour Integration, a retrospective longitudinal dataset that contains monthly calendar information about the labour states in the four years after completing vocational education. A wide range of states is covered: part-time employment, full-time employment, unemployment and several situations of inactivity. To examine young labour pathways we use sequence analysis, an exploratory technique which allows us to obtain a measure of dissimilarity between sequences. Later, a cluster analysis is applied that allow us to identify seven patterns of transition into the labour market. Nevertheless, most trajectories may be classified as linear, with a successful integration into the labour market. A multinomial logit model confirms that demographic and educational variables are relevant to explain the pathways followed by young people.

I. Introduction

School-to-work transition is no longer considered to take place at one point but it is seen as a sequential process with multiple stages that extend over a relatively long period in which individuals gradually settle down into the labour market (OECD, 2000; Ryan, 2001). Besides, routes between school and work are not always linear but involve breaks, divergences and reversals. According to this, school-to-work transition analysis has to take into account this transitory period as a whole, and it is not sufficient to focus on single events such as the moment to achieve a first job. On the contrary, single events have to be understood in their continuity and should not be isolated from each other. Approaching the entry process in this way, both the speed of finding jobs and the stability of jobs bring into focus. The speed has implications for the amount and duration of unemployment spells, while job stability approximates the quality of the match. Consequently, a more appropriate data to analyze the entry to labour market must contain the sequence of the labour states each individual is passing through to reach the labour market integration or, in other words, the serial succession of states describing the structure and composition of the labour pathway.

There is a large tradition of research arguing that the pace and the smoothness of transition to work are primarily determined by the educational background (Scherer, 2005; Corrales and Rodríguez, 2004). In this sense, it is interesting to know the role that vocational education plays in the labour market outcomes as well as in the way the transition between the two systems performs. Labour market entrants already qualified for an occupation do not have to be trained (at least at the same extent as school leavers with only general education) and therefore, they are expected to find a job that matches more successfully.

1
2
3 However, the occupational specificity of vocational education makes difficult to
4
5 adjust skill demands to the rapid changes into the labour market. Then, different
6
7 outcomes, in the case of individuals with vocational education say something
8
9 about the real qualifications and skills of graduates as well as the degree of
10
11 mismatch between the labour demand and supply in some areas of vocational
12
13 specialization.
14
15
16

17 Vocational training in Spain has been subjected to a substantial reform in the
18
19 last years and new tackles into employment has been established.¹ The current
20
21 regulated vocational training is integrated into the education system and it is
22
23 divided into two levels: lower and upper vocational training (Homs, 2009). Lower
24
25 vocational training is inside the upper secondary education while upper vocational
26
27 training is an alternative to university.
28
29
30

31 Both of them have a modular structure of variable duration which goes from
32
33 1300 to 2000 hours, organised in one or two years. There are two types of
34
35 modules: theoretical-practical modules, aimed at developing a series of vocational
36
37 capabilities previously defined in relation to jobs, and a practical training module,
38
39 known as the Training in the Workplace module, which is done in a real
40
41 workplace (rather than in the educational centre) generally through internships in
42
43 a company. Students are supervised by a school tutor and a workplace tutor who
44
45 develop a work plan for them. The internship represents 25 percent of the teaching
46
47 time of a cycle and it is also evaluated. This work-linked training module does not
48
49 reach the level of alternation of Germany's dual system, nor the programmes of
50
51 training in alternation with employment of the French system, but it has
52
53 contributed to overcoming the traditional isolation of training centres from
54
55
56
57
58
59
60

¹ Nowadays, Spanish Vocational Training is ruled by the Act on the General Organization of the Education System (LOGSE) passed in 1990.

1
2
3 businesses and it recognizes of the importance of on-the-job training as an aspect
4
5 of vocational training.
6
7

8 It is also important to point out that the vocational training cycles are
9
10 organized by occupational families, which are large groupings by occupation or
11
12 vocation, some of a more sector character and others more transversal but which
13
14 cover the majority of vocational content in the labour market. The National
15
16 Institute of Qualifications updates periodically the number and definition of these
17
18 occupational families. Nowadays, there are 26 occupational families, which cover
19
20 the primary sector, with farming activities and maritime and fishing activities; the
21
22 secondary sector, with for example, chemistry, mechanical production, electricity
23
24 and electronics, graphic arts and the tertiary sector, such as hotel industry and
25
26 tourism, socio-cultural and community services, health, etc.
27
28
29
30

31 With respect to lower vocational training, it constitutes the path for vocational
32
33 specialization for those students who successfully complete basic education² and
34
35 want to join to the labour market.³ It could be said that vocational training is a
36
37 basic means of providing directly access to employment. So, if the system fails to
38
39 match workers' occupational profiles to economic and social requirements,
40
41 serious consequences in terms of exclusion may appear.⁴
42
43
44
45

46 On the other hand, Spanish youth labour market has faced some changes
47
48 derived from growing flexibility, expansion of educational system and the skill
49
50 biased technological progress (Dolado *et al.*, 2000) that may also affect the
51
52
53
54

55 ² Access to the lower-level cycles can be attained once the certificate of basic education has been
56
57 obtained. This requirement, which insures a minimum level of knowledge and abilities, has
58
59 contributed to increase the quality of these training cycles.

60 ³ In the academic year 2000/2001 there were 304580 people enrolled in lower vocational training system, of which 130715 were women. This represents 28% of total students enrolled in upper secondary education and 24% of women. In 2008/2009 the rates have increased and 32% of all students enrolled in upper secondary education opt for vocational studies (28% of women).

⁴ Also, low retention rates and drop-outs could increase the exclusion problem.

1
2
3 transition to work for vocational students. These authors have provided evidence
4
5 of a crowding-out effect, what means that the lower educated workers are replaced
6
7 by higher educated ones in their traditional entry jobs.
8
9

10 The aim of this article is to identify typologies of transitions from school to
11
12 work in the early years after leaving vocational lower-secondary education in
13
14 Spain in the light of the last changes in vocational training. A detailed
15
16 examination of these patterns allows identifying the pathways which conduct
17
18 more frequently to a job as well as to detect which fail to integrate students into
19
20 the labour market. This information would be of great usefulness for designing
21
22 political measures aimed to facilitate the integration of young people into the
23
24 labour market and to design efficiently vocational programmes.
25
26
27
28

29 The analysis consists of two parts. First, the youth labour trajectories are
30
31 compared using the sequence analysis. This technique was first introduced to
32
33 Social Sciences in the form of Optimal Matching Analysis (OMA). The idea
34
35 behind OMA is that the similarity between sequences can be expressed as the cost
36
37 of turning one sequence into another by applying a set of operations (substitution,
38
39 insertion and deletion). We apply a recent extension of sequence analysis, the
40
41 modified Hamming dissimilarity measure (Lesnard, 2006) that enables us to
42
43 account for time dependent substitution costs. This initial step ends up by
44
45 applying a cluster analysis taking as input the distance matrix obtained previously,
46
47 and grouping trajectories that are very similar. As a second step, a multinomial
48
49 logit model is estimated to evaluate the role of some demographic and educational
50
51 variables in determining individual pathways.
52
53
54
55
56

57 The paper is organised as follows. The next section provides a short review of
58
59 literature on this topic. Section 3 gives a description of data used in the empirical
60

1
2
3 analysis. The fourth section explains details of the techniques applied (sequence
4 analysis and cluster analysis) and some specifications in the application to our
5 data. Next, the results of a multinomial logit analysis are presented and discussed.
6
7
8
9
10 A robustness analysis is included in section 6 and the final section incorporates
11 the main conclusions.
12
13
14

15 16 **II Review of literature**

17
18
19 The study of the first attempts to enter into the labour market has raised great
20 interest among labour economists and a large literature can be found on this issue
21 (OECD, 2000; Ryan, 2001; Raffae, 2003; Davia, 2004; Casquero, 2007 for the
22 Spanish case). One of the most relevant results is that the school-to-work
23 transition process has become longer and more complex due to the growing
24 importance of intermediate states such as temporary jobs, rapid job-changing and
25 instability. However, little research has examined each trajectory as a whole. In
26 particular, we only find few papers handling this subject, mainly with a cross-
27 national comparison approach (Scherer, 2001; McVicar and Anyadike-Danes,
28 2002; Brzinsky-Fay, 2007; Quintini and Manfredi, 2009). All of them apply a
29 more recent and specific methodology to measure similarity between trajectories,
30 the sequence analysis, and make homogeneous groups of trajectories using the
31 cluster analysis. Sequence analysis allows them to handle the whole trajectory
32 taking into account the information about the length and the ordering of states. In
33 Scherer (2001), British and German youth trajectories are compared in order to
34 discover distinct patterns of transition into the labour market and to know how
35 they differ between different institutional contexts. McVicar and Anyadike-Danes
36 (2002) focus on the transition process from education to labour market in Ireland.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The main goal for these authors is to identify those background characteristics that
4 have influence on the chances of experiencing a particular type of transition.
5
6 Brzinsky-Fay (2007) compares the school-to-work transitions of young people for
7
8 several European countries (included Spain) and classifies them in eight different
9
10 typologies trying to find differences across countries supported by the country
11
12 classification described in Gangl (2003).⁵ The dominant trajectory in Spain is the
13
14 failure type, where unemployment is the primary state.
15
16
17
18

19
20 The most recent application of sequence analysis can be found in Quintini
21
22 and Manfredi (2009) where school-to-work transition pathways in the United
23
24 States and Europe are compared. Pathways in the United States are characterised
25
26 by more dynamism than in Europe, although some similarities are found. Again,
27
28 the dominant trajectory in Spain is the failure one. Also, Spain has the largest
29
30 share of school-leavers in the “recovery” pathway, experienced prolonged
31
32 unemployment but ultimately settle into work.
33
34

35
36 However, these papers are not strictly focused on the school-to-work
37
38 transitions for people with vocational education. Studies focusing on the effect of
39
40 vocational education have a more sociological perspective (Dolton *et al.*, 1994;
41
42 Arum and Shavit, 1995; Bonnal *et al.*, 2002; Shavit and Muller, 2000; Ryan,
43
44 2003; Iannelli and Raffe, 2007). These studies do reach some general conclusion.
45
46 Education systems in which students spent time learning and working in an
47
48 enterprise, are generally noted for having low levels of youth unemployment and a
49
50 smooth transition to work, because vocational training provides a close link
51
52
53
54
55
56
57
58
59

60
⁵ Gangl (2003) distinguish between three patterns of school-to-work transitions depending on differences in educational systems. Spain is classified as a country with a low level of educational attainment and a weakly developed vocational training system.

1
2
3 between education and work and also it is thought to provide a specific form of
4 socialization to working life.⁶
5
6
7

8 Spanish research dealing with typologies of transition has basically followed
9 a more qualitative approach (Cachón, 2000). According to this research, the
10 predominant situation for young people in Spain is the “successive
11 approximation”, but the “precariousness” is actually playing a major role.
12 However, the corresponding for young people with vocational training is the
13 “precocious success”, as they do not wait a long time to enter employment and
14 their first jobs are quite stable.
15
16
17
18
19
20
21
22
23
24
25

26 **III. Data and basic concepts**

27
28 The data we analyzed is extracted from the Survey on Educational-Training
29 Transition and Labour Integration (ETEFIL-2005).⁷ This survey has been the
30 consequence of the concern about the high level of youth unemployment in Spain.
31 Its main goal is to provide the essential elements to study jointly the pathways
32 followed by young people both in the educational system and in the transition
33 from education to work.
34
35
36
37
38
39
40
41
42

43 The sample comprises young people (younger than 26 years old in 2001) who
44 finished non-university studies or abandoned the educational system without
45 completing compulsory secondary education in the 2000/01 academic year. For
46 the purpose of our analysis, we only deal with a subsample, the one formed by
47 young people who finished vocational lower-secondary education. It consists of
48
49
50
51
52
53
54
55
56
57
58
59
60
7612 cases.

⁶ To know more about several theories explaining the vocational effect on the transition from school to work see Iannelli and Raffe (2007).

⁷ See INE (2007) for a more detailed description of data.

1
2
3 The longitudinal structure of the survey, based on retrospective information
4 on what young people have been doing in the early years after completing
5 vocational education, allows us to know exactly the monthly serial succession of
6 states which describes the structure and composition of the labour pathway. The
7 pathway begins when individual leaves this level of education and ends when is
8 interviewed in 2005. Hence, the survey approximately reproduces the next four
9 years after leaving education.
10
11
12
13
14
15
16
17
18

19 This information jointly with some other variables related to personal and
20 education characteristics are used in the analysis. With respect to the labour states,
21 the survey distinguishes up to 13 as a result of several combinations of the usual
22 labour situations: employment, unemployment and inactivity. Such detailed
23 information makes difficult the application of the proposed methodology.
24 Therefore, it was necessary to cut the amount of labour states. Taking into account
25 the overlap and interaction of educational and labour market situations in the early
26 years after finishing vocational education, seems reasonable to distinguish various
27 inactivity situations depending on the link to education (not studying, studying in
28 the formal educational system or studying non-formal education). On the other
29 hand, we have also distinguished between part and full-time jobs (working less or
30 more than 20 hours a week).⁸ In short, the labour states contemplated appears in
31 Table 1 and are the basic elements from which the trajectories are constructed.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 [Table 1 about here]
51
52

53 Data on vocational training allow us to distinguish between gender and a
54 broad number of areas of specialization (occupational families). These broad
55 specialities include: agriculture, business, health care, trade, industry, marketing
56
57
58
59
60

⁸ In any case, employment has always been the prevalent state. For instance, those individuals who are working part-time and studying at the same time are considered to be in the employment state.

1
2
3 and services (see Table 2). It must be noted that there exists gender segregation by
4
5 specialities. Some professional families are dominated by men, for example
6
7 Electricity and electronics, whereas women are more prone to families related to
8
9 Services sector as Health or Sociocultural and community services.
10
11

12 [Table 2 about here]
13
14

15 16 **IV. Analysis of youth labour trajectories** 17

18 *A. Methodology* 19

20
21 The final objective is to classify the labour trajectories into groups, in such a way
22
23 that trajectories belonging to the same group will be very similar and
24
25 simultaneously, very different from trajectories in other groups. For this purpose,
26
27 sequence analysis is used to compare the entry to labour market for young people
28
29 with vocational lower-secondary education.
30
31
32

33
34 In the literature there are two approaches capable of systematically
35
36 describing the sequence of events and their duration (Robette and Thibault, 2008).
37
38 One of them is the Qualitative Harmonic Analysis (QHA), a factor analysis
39
40 method developed by French statisticians in the 1980s which takes account of
41
42 time (Espinasse, 1993; Robette and Thibault, 2008). The other is Optimal
43
44 Matching Analysis (OMA), a set of algorithmic techniques imported from the Life
45
46 sciences by American sociologists in the late 1980s which also emphasizes the
47
48 order of events.
49
50
51

52
53 Before explaining the basic issues related to the OMA and discuss some
54
55 methodological aspects of its application to our research problem we will describe
56
57 briefly the main characteristics of QHA in order to understand the differences
58
59 between these two methods.
60

1
2
3 Harmonic analysis is a branch of mathematics widely applied in Physics and
4
5
6 Biology. Its use in the Social Sciences is more recent, and dates back to the 1970s
7
8 (Deville, 1974). The aim at that time was to use the notion of duration to explain
9
10 social phenomena via data on individual trajectories.
11

12
13 QHA involves defining an observation period, dividing it into a finite
14
15 number of intervals then measuring for each individual the proportion of time
16
17 spent in each of the states in each interval.⁹ A correspondence analysis on the
18
19 matrix thus formed, with individuals in the rows and time spent in each state
20
21 during each interval in the columns, provides a means to summarize the
22
23 information by selecting the factors with the highest inertia (Deville, 1982). Then
24
25 using an ascending hierarchical classification is possible to determine typologies
26
27 of trajectories. The typologies thus constructed takes account not only of the
28
29 sequence of states, but also the periods spent in the different states. It also keeps
30
31 all individuals in the analysis, reducing data heterogeneity with minimum
32
33 information loss.
34
35
36
37

38
39 As Robette and Thibault (2008) remark harmonic analysis performs better
40
41 when attention is focused on the duration of certain stages and to distinguish small
42
43 and specific groups. By contrast, optimal matching is a better choice for analysing
44
45 trajectories by type of transition and when mobility is the focus of attention, given
46
47 that is more sensitive to transitions, notably at the start of the trajectory when
48
49 changes are more frequent. In our case, OMA is also considered to be more
50
51 appropriate as allow us to take into account the point in time at which
52
53 transformation is carried out by using time-dependent substitution costs.
54
55
56
57
58

59
60

⁹ As derived from explanation, QHA relies on two choices: the number of intervals in which the study period must be divided and the amplitude of intervals, because these amplitudes do not necessarily have to be of equal length.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

The Optimal Matching Analysis (OMA) was first introduced to the Social Sciences by Abbott and Forrest (1986) and was originally created for alignment of sequences in the Biological Sciences.¹⁰ As it was mentioned above, the idea behind OMA is that the similarity between sequences can be expressed as the cost of transforming one sequence into another by applying a set of operations. There are three types of operations: substitution (a state is substituted by another one), insertion (a state is inserted into the sequence) and deletion (a state is deleted altogether).¹¹ The latter two operations are subsumed to so-called indel operations.¹² As noted in the literature, substitutions emphasize the ordering of events whereas indel operations stress the occurrence of events. Each type of operation is assigned a cost. The OMA looks for the less “expensive” combination of operations required to transform one sequence into another and assume this cost is the distance that separates both sequences (the similarity distance). Then, a pair of sequences with a small distance means that they are similar, while pairs with a large distance are more distinct. Such comparison is repeated for all sequence pairs in the sample to obtain a distance matrix that measures how much every sequence resembles each other. This distance matrix is used in the cluster analysis to identify representative pathways by grouping together sequences that are similar to each other, i.e. that have the smallest distance between themselves.

Therefore, the first step in sequence analysis is to assign costs to each operation. Nevertheless, we focus in another question before beginning to explain how to assign costs to each operation: the length of trajectories. As it was

56
57
58
59
60

¹⁰ For a review, see Abbott and Tsay (2000).

¹¹ Substitution is also interpreted as a combination of an insertion and a deletion. For example, transforming ABA into AAA involves deleting B and then inserting an A.

¹² These two operations used to be treated in the same way, given that deleting an element from a sequence is the same as inserting this element in the other sequence. For example, the difference between sequences ABA and AA can be eliminated either by removing B from ABA or adding B to AA.

1
2
3 explained in the data section, survey allows us to collect information about the
4
5 labour states occupied between two dates: the date of completion vocational
6
7 education and the date of survey. But the month of completion vocational
8
9 education and the month of interview are not always the same for all individuals.
10
11 Consequently, the length of sequences is not the same. However, the measure
12
13 used here, the dynamic Hamming dissimilarity (Lesnard, 2006), can only handle
14
15 sequences of equal length. For this reason, youth labour trajectories have been
16
17 truncated and only contain the states occupied since vocational training is finished
18
19 for a further three years. In this way, sequence length is 36 months.
20
21
22
23
24

25 Coming to the issue of assignment costs, it must be said that this step is
26
27 critical because different costs have influence on the resulting distance matrix.
28
29 Ideally, setting values to operations costs should be based on theoretical grounds
30
31 but there is not always a theory supporting this. Given that in our context, there
32
33 are no accepted measures of quantitative differences between states to guide us in
34
35 the assignment of the costs (Malo and Muñoz-Bullón, 2003), we have followed
36
37 the criteria established by the dynamic Hamming dissimilarity measure (Lesnard,
38
39 2006), where the substitution costs, derived from the data itself, are time
40
41 dependent, and the indel costs have no sense.¹³ This strategy has also been applied
42
43 to study the employment mobility (Malo and Muñoz-Bullón, 2003) and the string
44
45 to adulthood (Aassve *et al.*, 2007).
46
47
48
49

50 Hence, the cost of substituting a state (for example, unemployment) by
51
52 another one (employment) is assumed that it is related to the transition frequencies
53
54 between those states. More specifically, substitution costs are inversely
55
56 proportional to frequencies, that is, the less frequent is a transition between two
57
58
59
60

¹³ Assign costs to indel operations is useless when the sequences have equal length.

1
2
3 states, the bigger is the cost to substitute one of the states by the other one.
4
5
6 Besides, we accept that substitution costs also vary with time. That means that
7
8 probability of staying in a state, for instance employment, two years later after
9
10 completing vocational training is not the same as to get a job just right away after
11
12 ending. In this sense, the transition matrix which contains the transition
13
14 frequencies between the six states is not fixed, but time dependent. Summarizing,
15
16 the substitution costs have been calculated using the monthly transition
17
18 frequencies derived from data taking into account the following expression,
19
20
21

$$22 \text{ cost } (a,b) = \begin{cases} 4 - p(X_t = a / X_{t-1} = b) + p(X_t = b / X_{t-1} = a) + \\ 23 p(X_{t+1} = a / X_t = b) + p(X_{t+1} = b / X_t = a) & a \neq b \\ 24 \\ 25 \\ 26 \\ 27 \quad \quad \quad 0 & a = b \\ 28 \end{cases}$$

29 where p stands for the transition frequencies between state a (or b) to state b (or a)
30
31 in a given month. Each matrix of substitution costs is symmetric and its size is
32
33 equal to the number of states.
34
35

36 The dynamic algorithm evaluates all possible solutions (combinations of
37
38 operations to transform a sequence into the other) and returns the “cheapest”
39
40 one.¹⁴ As a result, a symmetric matrix with size equal to the number of sequences
41
42 that contains the similarity between each pair of sequences is obtained. The matrix
43
44 is called a distance matrix as each cell (a_{ij}) contains the distance between the
45
46 individual sequence i and individual sequence j . The cells on the diagonal of the
47
48 matrix contain only zeros because they represent the distance between a sequence
49
50 and itself.
51
52
53
54
55
56
57
58
59

60 ¹⁴ We use a Stata program created by Lesnard (2006). The program generates several files that contains the series of monthly transition matrices, substitution cost matrices and the distance matrix.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Next, we apply a cluster analysis on this distance matrix to group sequences into an undetermined number of clusters which identify representative pathways. A difficulty to apply the cluster analysis is the great number of repeated sequences that appear in our data.¹⁵ As Martin *et al.* (2008) noticed cluster analysis is affected by ties. For this reason, the cluster analysis is only applied to unique sequences, that is, we eliminate the sequences repeated.

Cluster analysis is a well-known technique that requires making some decisions about the different clustering algorithms and the identification of the appropriate number of clusters. In relation to the cluster algorithm, there exists a large amount of methods of grouping and there is not a clear criterion to decide between them. We have chosen an agglomerative hierarchical method, the Ward method, because it finds most homogeneous clusters and it is the most popular between similar studies (Scherer, 2001; Brzinsky-Fay, 2007; Martin *et al.*, 2008). Besides, it seems to provide the most reasonable results when we compare the performance of different cluster algorithms.

Like other hierarchical agglomerative methods, at the start of the process, each cell in the distance matrix – each pair of sequences ij – represents a cluster and has a distance a_{ij} associated to it. These primitive clusters are then progressively merged and, at every round, the distance between each two clusters is recalculated using the criterion applied by the Ward linkage.

The Ward's clustering criterion is based on the error sum-of-squares, ESS, which is defined as the sum of the squared distances of individuals from the centre of gravity of the cluster to which they have been assigned. Initially, ESS is 0,

¹⁵ A previous description of sequences based on aggregated measures related to the number of states, the time spent on states and most frequent sequences has discovered that a high percentage of young people share the same trajectory.

since every individual is in a cluster of its own. At each stage the link created is the one that makes the smallest increase in the ESS.

In mathematical terms, the error sum-of-squares of a group r is given by

$$ESS_r = \sum_{m=1}^{n_r} \sum_{j=1}^p (X_{rjm} - \bar{X}_{rj})^2$$

where X_{rjm} denotes the value of variable X_j in the m element of group r (in our case the variables will be the distance between trajectories included in cluster r).

In each step the algorithm merges those groups r and s that minimize

$$ESS_t - (ESS_r + ESS_s) = \frac{n_r n_s}{n_r + n_s} d_{rs}^2$$

with t is the combine cluster resulting from the fusion cluster r and s and d_{rs}^2 the squared Euclidean distance between the center of gravity of clusters r y s .

The final question to be solved is the identification of the appropriate number of clusters. We have considered the change in the agglomeration coefficient represented in the dendrogram to take this decision.¹⁶ Nevertheless, after comparing some different results, the final solution (seven clusters) has also taken into account contextual arguments, an internal analysis of homogeneity and the characteristics of clusters that help to understand what youth in each cluster have in common and whether each cluster has analytically meaningful.¹⁷

B. Typologies of trajectories

We begin by describing the clusters and the characteristics of those youths belonging to each one.¹⁸ The mean number of months spent in each state by

¹⁶ Other authors propose to display and visualize results using self-organizing maps (Giret and Rousset, 2007).

¹⁷ To evaluate the robustness of the cluster analysis, we have investigated the presence of more or fewer clusters, but the identification described in the paper with seven clusters was the clearest.

¹⁸ To describe clusters we retrieve the sequences repeated.

1
2
3 members of each cluster and the mean number of episodes in each state are
4 included in Table 3. Bold fields indicate the dominant states within each cluster.
5
6 Some other background characteristics such as gender and professional family are
7
8 also considered to describe clusters (Table 4). Finally, some figures, representing
9
10 the monthly states distribution for each cluster, allow us to differentiate clusters
11
12 not only for their composition but for the location and order of states and give us a
13
14 more detailed picture of the transition process, making yet easier to recognize
15
16 which is the dominant state in each cluster (see Figure 1). A combination of all
17
18 these measures helps us to evaluate if the trajectories could be considered linear
19
20 transitions or non-linear transitions, having defined a non-linear transition as the
21
22 one that involves breaks, changes of direction or unusual sequences of events
23
24 (Furlong *et al.*, 2006).
25
26
27
28
29
30

31
32 **[Table 3 about here]**
33

34 The largest group by far is the “continuously working in a full-time job”.
35
36 More than one third of the young people (70%) have followed a trajectory
37
38 basically composed by the “full-time employment” state. Within this cluster, the
39
40 mean number of months in employment is 33, clearly superior to the number of
41
42 months for the total sample. Also, this cluster is characterised by the diversity of
43
44 sequences as it contains the highest number of different sequences (1350). Thus,
45
46 occasional and short spells in other states can be found dispersed across the
47
48 trajectory. Continuous full-time work participation is slightly more common
49
50 among males than females.
51
52
53

54
55 Cluster 2 (“part-time job”) is also characterized by sequences essentially
56
57 formed by employment, but in contrast to the first cluster, young people spent
58
59 most of the first three years after finishing vocational education working in part-
60

1
2
3 time jobs. Over the whole period (36 months), the mean duration in part-time
4 employment for members of this group was 24 months. Considering both the time
5 spent in part-time and full-time jobs, the total duration is similar to the previous
6 cluster. The participation in the labour market tends to be full-time at the end of
7 the trajectory as it could be appreciated in Figure 1. Approximately 3.5% of young
8 people belong to this cluster and it is clearly a female cluster (59%).
9
10
11
12
13
14
15
16

17 Sequences belonging to cluster 3 (“non-formal education”) are basically
18 composed of inactivity periods (studying in not formal education) and
19 employment periods (3.4% of individuals belongs to this cluster). In general,
20 people from this cluster distribute their time spending 18 months studying and just
21 over one year working in full-time jobs. A closer look at Figure 1 shows another
22 important explanation of how the trajectories are formed; inactivity periods
23 usually are part of the early years of the trajectory and employment periods are
24 more likely to appear at the end of the trajectory. This cluster hardly display
25 gender imbalance.
26
27
28
29
30
31
32
33
34
35
36
37

38 A typical sequence from cluster 4 (“full-inactivity”) corresponds to an
39 individual spending just about two thirds of time not studying and not working
40 (total inactivity). This cluster could be the most problematic given that the mean
41 duration in employment is the smallest one (6 months) and inactivity is a common
42 issue along the trajectory. Fortunately, not many young people belong to this
43 cluster (approximately 2% of the sample, mostly females).
44
45
46
47
48
49
50
51
52

53 There are two clusters (5 and 7) where the “inactivity but studying in formal
54 education” state is the dominant, but we can appreciate some important
55 differences between them. Young people in cluster 5 (“return to formal
56 education”) spend at least half of the time studying (in cluster 7 time spent in this
57
58
59
60

1
2
3 state is longer, 23 months) but what really differentiate them is that they return to
4 education after a period away. Besides, this cluster seems to be the more
5 heterogeneous taking into account the time spent in other states. The break or
6 change in the direction of the transition to work that we appreciate in this cluster,
7 it makes possible to describe the transition period as not linear, according to the
8 above definition. The cluster 7 (“continue formal education”) is the second largest
9 group (13%) and corresponds to young people that decide to improve their
10 education before entering into the labour market. Nevertheless, the size of these
11 two clusters is lower than that from cluster 1, which means that vocational lower-
12 secondary education is more likely to lead to employment and less likely to lead
13 to higher education.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 From a political and social point of view, sequences in cluster 6
30 (“unemployment”) could be considered those having the worst behaviour. There is
31 a clear dominance of unemployment, with an average of 24 months distributed in
32 two spells of unemployment. Long spells of unemployment are usually found at
33 the beginning of the trajectory. This type of trajectory is only shared by the 5% of
34 the total sample and it is clearly a female trajectory.
35
36
37
38
39
40
41
42

43 One of the main questions address by this study is whether qualifications and
44 skills of graduates have different effects in outcomes in the case of individuals
45 with vocational education. For this reason, we examined in detail the distribution
46 of the different occupational families by cluster (see Table 4). First of all, it
47 should be noted that there are some occupational families more numerous.
48 Students in vocational education tend to choose: Administration and management
49 (18.5%), Health (12.5%), Electricity and electronics (11.3%) and Transport and
50 maintenance of motor vehicles (8.6%). The trajectory described in cluster 1
51
52
53
54
55
56
57
58
59
60

1
2
3 continues being the most frequent even when we distinguish by professional
4 families. Nevertheless, the percentage of young people with this kind of trajectory
5 varies considerably, from 85.5% for Mechanical manufacturing to 55.8% to
6 Chemistry (or 32.1% in the case of Arts and craftwork). As it can be observed, the
7 occupational families occupying the first positions are not the same that present a
8 higher number of students. Also cluster 7 (“continue formal education”) is the
9 second trajectory more frequent for all families. In particular, those families for
10 which paths are more related to education, postponing the entry to labour market,
11 are Arts and Craftwork, Agriculture activities, Chemistry, Image and sound and
12 Hotel industry and Tourism.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 For some occupational families the partial-time job is more frequent than
28 unemployment. For example, the trajectory of part-time employment is more
29 common for the Chemistry, Image and sound and Construction and civil work
30 families. Surprisingly, a pathway basically composed of unemployment is not
31 found for those students who studied the Construction and civil work family. An
32 explanation to this result maybe the favourable economic situation for this sector
33 in the analised period.
34
35
36
37
38
39
40
41
42

43 [Table 4 about here]
44
45
46

47 **V. Multivariate analysis: determinants of belonging to clusters**

48
49

50 In order to provide a more comprehensive treatment of differences between
51 clusters we have estimated a multinomial logit model. The model allows us to
52 isolate the effect of some demographic and educational factors on the probability
53 of each individual to belong to a particular cluster. In this respect, we evaluate the
54 role of these variables in determining individual trajectories. The dependent
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
variable for the model takes values 1 to 7 and reflects the cluster to which individual is assigned. As explanatory factors we have included gender, age, the geographical area where the individual finished his studies and several dummies collecting the area of specialization in vocational education (occupational families). It has been necessary to join some occupational families in order to get enough size for each cluster.¹⁹ As mentioned above, young men and women do not have the same preferences when deciding to study vocational education and our data show clearly this gender segregation by specialization. Thus the colinearity between these two factors makes difficult to obtain accurate estimates when both of them are jointly included. For this reason, we have estimated two alternative specifications (each one with one of the two variables) to evaluate the impact of gender and occupational family in the probability of belonging to a cluster.

34
35
36
37
38
39
40
41
42
Before commenting the results, we have to note that the independence of irrelevant alternatives assumption, that is whether each of the dependent variable categories can be considered statistically independent of the others, has not been rejected (see Table 6).

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Estimation results of both specifications appear in Table 5 and are presented as relative risk ratios which, in the case of a quantitative variable, give the proportionate change in the relative risk of belonging to a given pathway rather than the full-employment trajectory when the variable changes by one unit. In the case, of a qualitative variable it gives the proportionate change in the relative risk for the collective of young people define by the variable compare with the reference category. For example, the relative risk ratio of 1.98 for the female

¹⁹ Information about the families that are aggregated could be found in Table 2.

1
2
3 variable in the part-time pathway outcome indicates that females have a relative
4 risk of belonging to the part-time pathway rather than full-time pathway 98%
5
6 higher than males.
7
8

9
10 Regarding gender the results show that females have a higher relative risk
11 ratio of belonging to part-time, full inactivity, unemployment and continue formal
12 education clusters rather than full-employment than men.
13
14

15
16 Turning to age, both specifications show that as age increases the relative
17 probability of belonging to clusters more related with education (cluster 3, 5 y 7)
18 against full-employment reduces.
19
20
21

22
23
24 [Table 5 and 6 about here]
25
26

27 There are also significant differences when considering the area of residence
28 as a proxy of the situation of local labour markets. The common results to both
29 specifications, indicates that the success in transitions after finishing vocational
30 education is greater if young is living in Madrid, because those young have a
31 lower risk ratio to follow any trajectory distinct to the full-time trajectory (the
32 reference area is the south of Spain which is characterized by the largest
33 unemployment rate). Results also suggest that living in eastern and north-eastern
34 regions also reduces the risk ratio of belonging to unemployment cluster, that is,
35 there is advantaged in terms of the availability of jobs in these local labour
36 market.
37
38
39
40
41
42
43
44
45
46
47
48
49

50 The area of specialization chosen has also significant explanatory power for
51 the type of pathway followed. If the occupational family is Agriculture, Services1,
52 Services2, Chemistry or Health the relative risk ratio to belong to cluster 7
53 (continue formal education) is higher. On the contrary, Industry and Construction
54 specialization reduces this risk ratio and also the risk ratio to belong to any cluster
55
56
57
58
59
60

1
2
3 distinct to the full-time employment. Health, Chemistry, Services1 and Services2
4
5 are also specializations where a part-time pathway has greater probability than
6
7 full-time pathway, and having studied Electricity and electronics and Transport
8
9 reduce the risk ratio of belonging to unemployment cluster.
10
11

12 Behind these results is the productive model prevailing in Spain in those
13
14 years of economic expansion, based on the construction sector, its ancillary
15
16 industries and on domestic demand. As a result there was a growing supply of
17
18 semi-skilled jobs in those sectors that could be covered by these youth. Hence, the
19
20 success of the occupational families related to construction, industry, electronics
21
22 and electricity and transport.
23
24
25

26
27 On the other hand, the activities most closely related to the services sector
28
29 and health have also had a development and generated jobs, but in these sectors
30
31 there has been much more competition among the youth with vocational training
32
33 and university studies that eventually has led to the need of more education, in
34
35 some cases, combine with part-time jobs.²⁰
36
37
38
39

40 VI. Robustness analysis

41
42 As already mentioned in the methodology, results depend greatly on the strategy
43
44 used to measure distances between sequences and the criteria to group sequences.
45
46 For this reason, we have considered two alternative specifications of the
47
48 substitution costs matrix to assess the robustness of the results. The first one
49
50 consists of setting to one the costs of substituting each pair of states, a usual
51
52 starting point substitution costs (Lesnard, 2010). As a second alternative, costs
53
54
55
56
57
58

59 ²⁰ At this respect, there exists a vast empirical literature that shows that Spain has suffered a
60 problem of over-education, mainly in the services sector, involving many young people with a
university degree in jobs which require less qualifications (Dolado *et al.*, 2000; Budria and Moro-
Egido, 2008).

1
2
3 have been theoretically grounded reflecting our own assumptions about the
4
5 inherent distinctions between the states. Given that the school-to-work-transition
6
7 process connects the educational system and the labour system, transitions
8
9 between states belonging to the educational system (or labour system) are
10
11 supposed to be easier and it is assigned lower costs than transitions from one
12
13 system to the other one. The transitions from/to full inactivity (nor studying nor
14
15 working) have been considered the most difficult one and it has the greatest costs
16
17 (see matrix in Figure 2). For each alternative, we have running a cluster analysis
18
19 to make groups of trajectories, taking as input the distance matrix obtained from
20
21 the optimal matching analysis with the corresponding substitution costs matrix. In
22
23 order to make easier the comparison with the previous results, the number of
24
25 clusters is the same as in the case of dependent substitution costs.
26
27
28
29
30
31

32 The degree of similarity of both alternatives in relation to the previous one
33
34 can be illustrated by a correspondence matrix which cross-tabulates the population
35
36 distributions obtained with two different partitions (Robette and Thilbaut, 2008).
37
38 It shows whether the individuals of a cluster from the first partition are
39
40 concentrated in a single cluster of the second partition or scattered between
41
42 several different ones. In Table 7, the seven-cluster distribution with time-
43
44 dependent substitution costs spans the rows and the seven-cluster distribution with
45
46 unit costs spans the columns, whereas at the bottom of this table, the seven-cluster
47
48 distribution with fixed costs spans the columns. In both cases, it is showed only
49
50 slight differences with the reference time-dependent costs chosen in previous
51
52 section, above all in the case of fixed costs.²¹ Most sequences are grouped in the
53
54 same clusters, what supposes that more than 70% of them are identical classified.
55
56
57
58
59

60

²¹ The robustness of the results with respect to different costs matrices was expected. Most of papers about the school-to-work-transition commented previously have also experimented with different variants of substitution costs and arrived at the same results.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Where we found more diversion is in those clusters related to inactivity as a consequence of studying in the formal education system. A detailed revision of the time-dependent substitution costs matrix confirms us that there are some similarities to the fixed theoretic matrix proposed here.

Also, a correspondence rate summarizing the degree of similarity between the two partitions can be calculated by means of the index proposed in Robette and Thibault (2008), which sum up the mode of each line and each column, dividing it by the double of the total sample size, that is

$$\sigma = \frac{\sum_i \max_j n_{ij} + \sum_j \max_i n_{ij}}{2N}$$

where n_{ij} represents the number of trajectories belonging simultaneously to the i th cluster of the first partition and the j th cluster of the second partition, and where N is the total population.

The correspondence between the distributions obtained with time-dependent costs and unit costs is 87% and between time-dependent cost and fixed costs 90%. The two alternative costs give similar results, a good indicator of their robustness.

VII. Conclusions

It has been widely documented that youth transitions have become increasingly long and complex what implies that few individuals follow identical routes to enter into the labour market and most faced routes involving movements in both directions (in and out of employment). However, the analysis of the school-to-work transition for Spanish young people with vocational lower-secondary education shows that pathways are not substantial distinct across individuals and most trajectories may be classified as linear, with a successful integration into the

1
2
3 labour market. This result is in agreement with educational systems including a
4
5 substantial training component in a real working context, in which the school-to-
6
7 work transition seems to be direct and smooth.
8
9

10 Specifically, the cluster analysis based on dissimilarity measures has
11 suggested seven coherent typologies of transitions from school to work.
12 Differences between them are basically due to the dominant labour state and to the
13
14 order and location of the states. The largest cluster by far is composed of
15
16 sequences characterized by a direct entry into the labour market with very short
17
18 spells in the other states. It is remarkable to note that this cluster is also the most
19
20 frequent when we distinguish by occupational families. This finding supports the
21
22 idea that the areas covered by the actual occupational families respond to the
23
24 current demand of labour market, at least, in this period in which the boom
25
26 experienced by the Spanish economy has made possible the integration of most
27
28 young people. The second cluster involves extended participation in education.
29
30 Again, this result seems to confirm an adequate behaviour of vocational students.
31
32 They try to improve their skills in order to get a better position into the labour
33
34 market. Those clusters more problematic, where young people experience
35
36 unfavourable long unemployment spells or spend most time without studying nor
37
38 working, are less numerous. In spite of its small size, it would be useful to take a
39
40 closer look at the characteristics of young people in the “full inactivity” cluster to
41
42 understand better their transition to work. But this requires further research.
43
44
45
46
47
48
49
50
51
52

53 On the other hand, differences between men and women have been found.
54
55 Women are more likely to experience trajectories with a greater participation in
56
57 part-time jobs or inactivity. In general, women sequences show greater variation
58
59 and its integration is less direct.
60

1
2
3 As a final thought, the paper highlights one point often overlooked when
4 transition to school-to-work is concentrated in a comparative analysis between
5 countries or in a national aggregated analysis. Spain is included in most European
6 studies as a country where young people experience serious difficulties to enter
7 into the labour market. Besides, some Spanish studies supported this statement.
8 However, this paper shows the opposite in the case of students with vocational
9 education.
10
11
12
13
14
15
16
17
18

19
20 The favorable situation of the Spanish economy has contributed greatly to the
21 rapid entry to the labour market. In this sense, it would be interesting to know
22 how the trajectories would be in the case of an adverse economic situation.
23 Besides, several aspects of the school-to-work transition (such as the quality of
24 jobs attained by these young people with vocational lower-secondary education)
25 have not been taken into account in this paper and they should help to improve our
26 analysis in the future.
27
28
29
30
31
32
33
34
35
36

37 **References**

- 38
39
40
41 Aassve, A., Billari, F. C. and Piccarreta, R. (2007) Strings of adulthood: a
42 sequence analysis of young British women's work-family trajectories,
43 *European Journal of Population*, **23**, 369-388.
44
45
46
47
48 Abbott, A. and Forrest, J. (1986) Optimal matching methods for historical
49 sequences, *Journal of Interdisciplinary History*, **16**, 471-494.
50
51
52
53 Abbott, A. and Tsay, A. (2000) Sequence analysis and optimal matching methods
54 in sociology: review and prospect, *Sociological Methods and Research*, **29**, 3-
55
56
57
58
59
60

- 1
2
3 Armstrong, D. and McVicar, D. (2000) Value added in further education and
4 vocational training in Northern Ireland, *Applied Economics*, **32**, 1727-1736.
5
6
7
8 Arum, R. and Shavit, Y. (1995) Secondary vocational education and the transition
9 from school to work, *Sociology of Education*, **68**, 187-204.
10
11
12
13 Bonnal, L., Mendes, S. and Sofer, C. (2002) School-to-work transition:
14 apprenticeship versus vocational school in France, *International Journal of*
15 *Manpower*, **23**, 426-442.
16
17
18
19
20 Brzinsky-Fay, C. (2007) Lost in transition? Labour market entry sequences of
21 school leavers in Europe, *European Sociological Review*, **23**, 409-422.
22
23
24
25 Brzinsky-Fay, C., Kohler, U. and Luniak, M. (2006) Sequence analysis with Stata,
26 *The Stata Journal*, **6**, 435-460.
27
28
29
30 Budria, S. and Moro-Egido, A. I. (2008) Education, educational mismatch and
31 wage inequality: evidence for Spain, *Economics of Education Review*, **27**, 332-
32 341.
33
34
35
36 Cachon, L. (2000) "National Report Spain", in *Documentation of European*
37 *Conference "Transition from initial vocational training into stable*
38 *employment"*, INBAS, Frankfurt am Main, 89-106.
39
40
41
42
43 Casquero, A. (2007) *Transición del sistema educativo al mercado laboral en*
44 *España. Especial referencia a la formación profesional*, Analistas Económicos
45 de Andalucía, Málaga (España).
46
47
48
49
50
51 Chan, T. W. (1995) Optimal matching analysis: a methodological note on
52 studying career mobility, *Work and Occupation*, **22**, 467-490.
53
54
55
56 Corrales, H. and Rodríguez, B. (2004) Transition from education to first
57 significant job in Spain: the influence of educational attainment, in
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Proceedings of the 58th International Atlantic Economic Conference, Chicago, Illinois.

Davia, M. A. (2004) *La inserción laboral de los jóvenes en la Unión Europea*, Consejo Económico y Social, Madrid.

Deville, J. C. (1974) Méthodes statistiques et numériques de l'analyse harmonique, *Annales de l'INSEE*, **15**, 3-101.

Deville, J. C. (1982) Analyses de données chronologiques qualitatives: comment analyser des calendriers ? *Annales de l'INSEE*, **45**, 45-104.

Dolado, J. J., Felgueroso, F. and Jimeno, J. F. (2000) Youth labour markets in Spain: education, training and crowding-out, *European Economic Review*, **44**, 943-958.

Dolton, P. J., Makepeace, G. H. and Trebble, J. G. (1994) The youth training scheme and the school-to-work transition, *Oxford Economic Papers*, **46**, 629-657.

Espinasse, J. M. (1993) Enquêtes de cheminement, chronogrammes et classification automatique, *Notes du Lirhe*, **159**, 93-19, CEREQ.

Everitt, B. S., Landau, S. and Leese, M. (2001) *Cluster analysis*, Oxford University Press.

Furlong, A., Cartmel, F. and Biggart, A. (2006) Choice biographies and transitional linearity: re-conceptualising modern youth transitions, *Papers*, **79**, 225-239.

Gangl, M. (2003) The structure of labour market entry in Europe: a typological analysis, in Muller, W. and Gangl, M. (eds.) *Transition from education to work in Europe*, Oxford University Press.

- 1
2
3 Giret, J. and Rousset, P. (2007) Classifying qualitative time series with SOM: the
4
5 typology of career paths in France, in *Proceedings of the 9th International*
6
7 *Work-Conference on Artificial Neural Networks*, Springer, Berlin, pp. 757-764.
8
9
10 Halpin, B. (2010) Optimal matching analysis and life-course data: the importance
11
12 of duration, *Sociological Methods and Research*, **38**, 365-388.
13
14
15 Halpin, B. and Chan, T. W. (1998) Class careers as sequences: an optimal
16
17 matching analysis of work-life histories, *European Sociological Review*, **14**,
18
19 111-130.
20
21
22 Homs, O. (2009): *Vocational training in Spain. Toward the knowledge of society*,
23
24 Social Studies Collection No. 25, Fundación La Caixa, Barcelona.
25
26
27 Iannelli, C. and Raffe, D. (2007) Vocational upper-secondary education and the
28
29 transition from school, *European Sociological Review*, **23**, 49-63.
30
31
32 Instituto Nacional de Estadística (INE) (2007) Encuesta de transición educativo
33
34 formativa e inserción laboral. Resumen Metodológico. Available at
35
36 www.ine.es/daco/daco42/etefil/notaetefil.pdf (accessed 17 December 2009).
37
38
39 Lesnard, L. (2006) Optimal matching and the social sciences, Working Paper No.
40
41 1/2006, CREST-INSEE, France.
42
43
44 Lesnard, L. (2010) Setting cost in optimal matching to uncover contemporaneous
45
46 socio-temporal patterns, *Sociological Methods and Research*, **38**, 389-419.
47
48
49 Malo, M. and Muñoz-Bullón, F. (2003) Employment status mobility from a life-
50
51 cycle perspective: a sequence analysis of work histories in the BHPS,
52
53 *Demographic Research*, **9**, 119-162.
54
55
56 Martin, P., Schoon, I. and Ross, A. (2008) Beyond transitions: applying Optimal
57
58 Matching analysis to life course research, *International Journal of Social*
59
60 *Research Methodology*, **11**, 179-199.

- 1
2
3 McVicar, D. and Anyadike-Danes, M. (2002) Predicting successful and
4
5 unsuccessful transitions from school to work by using sequence methods,
6
7 *Journal of the Royal Statistical Society (Series A)*, **165**, 317-334.
8
9
- 10 Merino, R. (2006) Two or three vocational training pathways? An assessment and
11
12 the current situation in Spain, *European Journal of Vocational Training*, **37**,
13
14 52-67.
15
16
- 17 OECD (2000) *From initial education to working life. Making transitions work*,
18
19 Paris: OECD.
20
21
- 22 Quintini, G. and Manfredi, T. (2009) Going separate ways? School-to-work
23
24 transitions in the United States and Europe, OECD Social, Working Paper in
25
26 Employment and Migration No. **90**, OECD, France.
27
28
- 29 Raffe, D. (2003) Pathways linking education and work: a review of concepts,
30
31 research, and policy debates, *Journal of Youth Studies*, **6**, 3-19.
32
33
- 34 Robette, N. and Thibault, N. (2008) Comparing qualitative harmonic analysis and
35
36 optimal matching: an exploratory study of occupational trajectories,
37
38 *Population-E*, **63**, 533-556.
39
40
- 41 Ryan, P. (2001) The school-to-work transition: a cross-national perspective,
42
43 *Journal of Economic Literature*, **39**, 34-92.
44
45
- 46 Ryan, P. (2003) Evaluating vocationalism, *European Journal of Education*, **38**,
47
48 147-162.
49
50
- 51 Scherer, S. (2001) Early career patterns: a comparison of Great Britain and West
52
53 Germany, *European Sociological Review*, **17**, 119-144.
54
55
- 56 Scherer, S. (2005) Patterns of labour market entry – long wait or career
57
58 instability? An empirical comparison of Italy, Great Britain and West
59
60 Germany, *European Sociological Review*, **21**, 427-440.

- 1
2
3 Schoon, I., McCulloch, A., Joshi, H. E., Wiggins, D. and Bynner, J. (2001)
4
5 Transitions from school to work in a changing social context, *Young*, **9**, 4-22.
6
7
8 Shavit, Y. and Muller, W. (2000) Vocational secondary education: where
9
10 diversion and where safety net? *European Societies*, **2**, 29-50.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Table 1. Classification of labour states

Code	Description
INEFSE	Inactivity but studying non-formal education
INESE	Inactivity but studying in the educational system
INNE	Inactivity and not studying
U	Unemployment
PT	Part-time employment (less than 20 hours per week)
FT	Full-time employment (equal or greater than 20 hours per week)

Table 2. Occupational families

Code	Description	Aggregation grouping
ADM	Administration and Management	Administration and Management
AGR	Agriculture activities	Agriculture
ART	Arts and Craftwork	Services2
CHE	Chemistry	Chemistry
CON	Construction and Civil Work	Industry and Construction
ELE	Electricity and Electronics	Electricity and Electronics
FOI	Food Industry	Industry and Construction
GRA	Graphic Arts	Services2
HEA	Health	Health
HOT	Hotel Industry and Tourism	Services1
IMA	Image and Sound	Services2
INS	Installation and Maintenance	Industry and Construction
MAR	Maritime Industry and Fisheries	Agriculture
MEC	Mechanical Manufacturing	Industry and Construction
PER	Personal Image	Services2
SPO	Physical and Sporting Activities	Services2
SOC	Sociocultural and Community Services	Services1
TEX	Textiles, Leather and Fur	Industry and Construction
TRM	Trade and Marketing	Services1
TRA	Transport and Maintenance of Motor Vehicles	Transport and Maintenance of Motor Vehicles
WFC	Wood, Furniture and Cork	Industry and Construction

Table 3. Description of clusters by the mean number spent in each state and the mean number of spells

Cluster	Title	N	INEFSE		INESE		INNE		U		PT		FT		
			% female	Dur.	Spell	Dur.	Spell	Dur.	Spell	Dur.	Spell	Dur.	Spell	Dur.	Spell
1	<i>Continuously in full-time job</i>	5373	42.8	0.22	0.04	0.18	0.03	0.45	0.10	1.91	0.54	0.23	0.04	33.01	1.37
				(1.40)	(0.22)	(1.21)	(0.23)	(1.97)	(0.37)	(3.61)	(0.93)	(1.45)	(0.22)	(4.69)	(0.80)
2	<i>Part-time job</i>	264	58.7	0.57	0.10	1.25	0.28	0.49	0.13	1.45	0.38	23.66	1.34	8.57	1.17
				(2.15)	(0.33)	(3.32)	(0.68)	(1.77)	(0.42)	(3.50)	(0.77)	(8.71)	(0.80)	(7.83)	(1.02)
3	<i>Non-formal education</i>	261	46.7	18.41	1.28	0.89	0.11	0.51	0.16	3.72	0.77	0.38	0.08	12.09	1.38
				(9.33)	(0.59)	(3.06)	(0.33)	(1.70)	(0.51)	(5.72)	(1.04)	(1.62)	(0.27)	(9.11)	(1.03)
4	<i>Full inactivity</i>	132	57.6	1.02	0.17	2.27	0.23	24.20	1.47	1.76	0.36	0.82	0.12	5.93	1.16
				(2.97)	(0.43)	(4.56)	(0.48)	(7.93)	(0.89)	(3.91)	(0.68)	(3.13)	(0.41)	(5.72)	(1.22)
5	<i>Return to formal education</i>	206	45.1	5.45	0.50	18.35	1.27	1.95	0.25	3.39	0.4	0.72	0.14	6.14	1.10
				(7.38)	(0.65)	(5.42)	(0.59)	(4.56)	(0.49)	(6.37)	(0.87)	(2.45)	(0.46)	(6.53)	(1.18)
6	<i>Unemployment</i>	383	66.8	0.48	0.12	0.19	0.03	0.37	0.07	24.01	2.09	0.63	0.12	10.33	1.82
				(1.57)	(0.39)	(1.25)	(0.20)	(1.82)	(0.29)	(7.47)	(1.26)	(2.46)	(0.41)	(7.16)	(1.44)
7	<i>Continue formal education</i>	993	47.0	0.37	0.06	23.09	1.30	0.24	0.06	1.60	0.30	0.69	0.12	10.00	1.11
				(1.71)	(0.25)	(9.58)	(0.69)	(1.23)	(0.26)	(3.92)	(0.63)	(2.77)	(0.43)	(9.38)	(1.06)
		7612	45.6	<i>1.05</i>		<i>3.76</i>		<i>0.87</i>		<i>3.06</i>		<i>1.15</i>		<i>26.11</i>	

Source: ETEFIL-2005, own calculations.

Notes: INEFSE: Inactivity but studying non-formal education, INESE: Inactivity but studying in the educational system, INNE: Inactivity and not studying, U: Unemployment, PT: Part-time employment, FT: Full-time employment.

Standard deviations are given in parentheses.

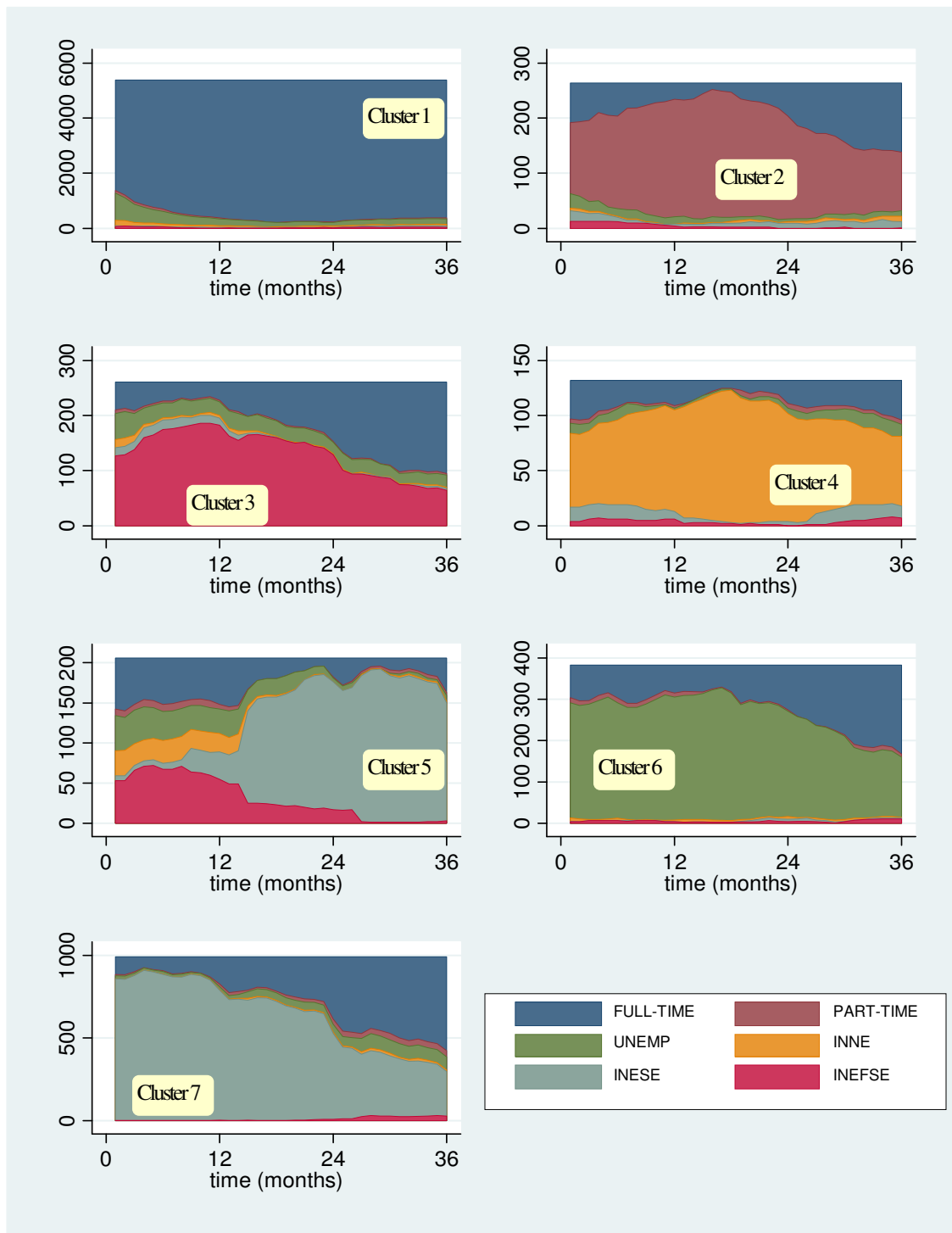
Table 4. Distribution of young people belonging to an occupational family by clusters (%)

Occupational family	Continuously in full-time job	Part-time job	Non-formal education	Full inactivity	Return to formal education	Unemployment	Continue formal education	Total
ADM	70.3	3.0	3.8	2.5	3.3	6.8	10.4	100
AGR	63.2	3.8	2.4	1.4	3.4	2.9	23.0	100
ART	32.1	3.3	5.7	1.6	2.9	6.1	48.4	100
CHE	55.8	8.5	3.1	1.8	3.1	7.6	20.1	100
CON	73.9	6.5	2.2	0.0	2.2	0.0	15.2	100
ELE	79.7	1.1	3.3	1.2	3.5	2.2	9.1	100
FOI	73.5	3.7	3.1	1.9	1.9	3.7	12.4	100
GRA	66.8	4.9	7.6	0.9	2.7	4.0	13.0	100
HEA	67.1	4.5	3.3	2.1	2.7	7.3	13.0	100
HOT	68.2	4.1	3.5	1.5	2.0	2.9	17.8	100
IMA	55.3	7.1	7.6	1.5	4.6	5.6	18.3	100
INS	80.6	1.3	1.9	0.6	3.4	2.2	10.0	100
MAR	60.0	12.0	4.0	4.0	0.0	0.0	20.0	100
MEC	85.5	1.0	1.8	1.3	2.1	2.8	5.4	100
PER	73.1	4.5	3.4	2.2	1.4	5.2	10.3	100
SOC	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100
SPO	57.0	6.1	3.6	2.4	4.9	7.3	18.8	100
TEX	62.3	4.4	1.5	1.5	2.9	13.0	14.5	100
TRA	83.2	2.3	2.0	0.8	1.2	1.8	8.7	100
TRM	63.1	5.0	4.3	2.8	2.2	8.2	14.5	100
WFC	82.6	2.4	1.9	0.9	1.9	5.7	4.7	100
	70.6	3.5	3.4	1.7	2.7	5.0	13.1	100

Source: ETEFIL-2005, own calculations.

Note: See Table 2 for occupational families.

Figure 1. Chronograms of the seven-cluster solution (number of individuals in each state)



Source: ETEFIL-2005, own calculations.

Note: INEFSE: Inactivity but studying non-formal education, INESE: Inactivity but studying in the educational system, INNE: Inactivity and not studying.

Table 5. Logit multinomial for the seven-cluster solution (relative risk ratios)

	Model 1						Model 2					
	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Female	1.98***	1.21	1.84***	1.15	2.76***	1.23**						
Age	1.03	0.91*	0.89	0.83***	1.01	0.94*	1.00	0.89*	0.88*	0.82***	1.00	0.92***
East Center	1.19 0.55*	0.75 0.83	1.06 1.28	2.07** 1.15	0.46*** 0.92	1.00 0.81	1.28 0.63	0.79 0.94	1.09 1.39	2.11** 1.23	0.47*** 1.01	1.07 0.90
Northeast Northwest	1.02 0.56	0.83 1.04	0.64 1.82	1.51 1.02	0.36*** 1.32	0.62*** 0.61**	1.17 0.66	0.98 1.22	0.71 2.05*	1.68 1.14	0.38*** 1.46*	0.72* 0.71*
Madrid	0.44***	0.56**	0.64	0.47**	0.34***	0.35***	0.53**	0.60**	0.68	0.48**	0.36***	0.40***
Agriculture Industry and Construction							1.53 0.55*	0.68 0.42***	0.70 0.35**	0.90 0.57*	0.33* 0.43***	2.03*** 0.66**
Services1 Services2							1.61* 1.87**	1.1 1.62*	0.98 0.87	0.71 1.05	0.90 0.88	1.70*** 2.27***
Transport Electricity and Electronics							0.67 0.32**	0.43** 0.78	0.26** 0.44*	0.34** 1.07	0.22*** 0.30***	0.73 0.83
Chemistry Health							3.12*** 1.64*	1.01 0.93	0.91 0.96	1.18 1.02	1.16 1.15	2.18*** 1.43**
Log L	-7837.7						-7719.1					
Pseudo R2	0.03						0.05					

Notes: Reference pathway is full-time job and individual reference is a female, with Administration studies and living in the Southern region. Standard errors are calculated and ***, **, * stand for statistically significant at 1%, 5% and 10% levels, respectively. See Table 2 for a description of occupational families groups.

Table 6. Small-Hsiao tests of IIA assumption

Omitted	Model 1			Model 2		
	chi2	P>chi2	evidence	chi2	P>chi2	evidence
Part-time job	28.990	0.901	for Ho	56.657	0.944	for Ho
Non-formal education	26.136	0.955	for Ho	63.622	0.823	for Ho
Full Inactivity	28.824	0.905	for Ho	65.808	0.767	for Ho
Return to formal education	30.016	0.875	for Ho	57.969	0.927	for Ho
Unemployment	29.215	0.896	for Ho	62.903	0.839	for Ho
Continue formal education	26.371	0.952	for Ho	63.276	0.831	for Ho
Full time job	31.375	0.833	for Ho	62.173	0.855	for Ho

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Figure 2. Fixed substitution costs

	U	FT	INEFSE	INNE	PT	INESE
U	0					
FT	1,2	0				
INEFSE	1,2	1,3	0			
INNE	1,4	1,4	1,4	0		
PT	1,2	1,1	1,3	1,4	0	
INESE	1,2	1,3	1,1	1,4	1,3	0

Note: See Table 1 for codes of states.

Table 7. Comparison of clusters with different substitution costs matrix

		<i>Unit costs</i>							Total
		1	2	3	4	5	6	7	
<i>Time-dependent substitution costs</i>	Full-time job	4970	9	362	2	1	18	11	5373
	Part-time job	7	249	0	1	0	1	6	264
	Non-formal education	129	0	21	91	1	16	3	261
	Full inactivity	9	4	2	112	2	3	0	132
	Return to formal education	0	0	5	26	153	22	0	206
	Unemployment	26	0	48	0	0	308	1	383
	Continue formal education	55	19	0	1	377	5	536	993
	Total	5196	281	438	233	534	373	557	7612
		<i>Fixed costs</i>							Total
		1	2	3	4	5	6	7	
<i>Time-dependent substitution costs</i>	Full-time	5292	8	4	0	1	57	11	5373
	Part-time	4	258	0	0	1	1	0	264
	Non-formal education	42	0	185	0	0	31	3	261
	Full inactivity	27	0	0	89	13	3	0	132
	Return to formal education	17	0	29	0	138	22	0	206
	Unemployment	85	0	0	1	0	297	0	383
	Continue formal education	79	19	2	0	372	27	494	993
	Total	5546	285	220	90	525	438	508	7612