Artículo de revista:

ORTÍZ-ÁVILA, Elsa; DEVOLDER, Daniel (2022) "Interactions between educational lifecycle and transition to adulthood: A proposal for a new questionnaire". *Advances in Life Course Research*, 52 (100465): 1-10 (ISSN: 1040-2608). https://doi.org/10.1016/j.alcr.2022.100465

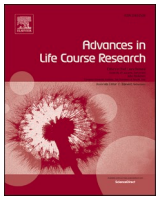
Advances in Life Course Research 52 (2022) 100465



Contents lists available at ScienceDirect

Advances in Life Course Research

journal homepage: www.elsevier.com/locate/alcr





Interactions between educational lifecycle and transition to adulthood: A proposal for a new questionnaire

Elsa Ortiz-Ávila [a,1], Daniel Devolder [b,*,2]

[a] Institute of Social Science and Human Studies, Autonomous University of the State of Hidalgo, Pachuca de Soto, Mexico
[b] Centre for Demographic Studies, Autonomous University of Barcelona, Spain

ARTICLE INFO

Keywords:
Education
Surveys
Transition to adulthood
Survival analysis
Questionnaire

ABSTRACT

Analysts who research the effect of education on behaviour related to the transition to adulthood are often limited by information available from censuses and surveys, which generally collect only the highest level of education attained at the time of the interview. Unfortunately, this leads to extreme simplification of people's educational paths and does not allow researchers to know exactly when they completed their studies, whether they did so satisfactorily, or whether they had periods of interruption in their educational cycle. In this paper we therefore investigate the problems deriving from having incomplete educational information when analysing events of transition to adulthood such as leaving the parental home, forming a first union, and having a first child. These needs for more detailed information explain why some surveys have questionnaires that would allow for reconstruction of the complete educational history, but this information is difficult to collect and to use. In order to overcome these difficulties of either the complexity or excessive simplicity of questionnaires on education, we propose a new set of questions. These would permit researchers to provide, with the minimum number of questions, the maximum possible useful information on the level and duration of studies, while also taking into account the most important and relevant interruptions. For the empirical analysis we have used the Fertility and Family Surveys (FFS) for 12 European countries with sufficient data to reconstruct the complete educational cycle.

1. Introduction

Educational level is an important factor in the explanation of differences in behaviour between individuals and consequently information on this variable is routinely collected in surveys or population censuses. In this study, we argue that the data obtained is generally too simplistic, or when it is detailed and complete as happens in a few surveys, it is too complex to collect and process. This can be especially problematic when analysing social processes that intermingle with periods of educational enrolment, for example transition to adulthood and the start of family life. We explain why this is so and propose a minimal set of questions that could be included in the questionnaire of future surveys collecting retrospective information on the lifecycle of individuals. The information on education which could thus be obtained presents the double advantage of being easy to process and analyse as well as being useful for investigation of the interactions with life-cycle events related with family formation.

In studies that take into account the effect of education on behaviour, the most commonly used indicator is the highest level of studies successfully completed. In our view, this may lead to oversimplification and even bias, as the information collected does not always allow researchers to know with certainty at what age individuals reached this level, whether they studied afterwards without obtaining another degree, or whether they interrupted their studies before reaching this final level. Scarcity of details on the educational career is normally due to the data sources used, most of which only collect basic information on studies. But this maximum level is the indicator used by analysts even if the source allows reconstruction of the educational history. This

* Corresponding author.
E-mail address: ddevolder@ced.uab.es (D. Devolder).

[1] Elsa Ortíz-Ávila acknowledges the help of the Public Secretary of Education through the research project PRODEP "La relación entre la finalización de los estudios y el inicio de la formación familiar en México".
[2] Daniel Devolder acknowledges the help of the Ministry of Economy and Competitiveness through the research contract on "¿Las personas mayores tendrán parientes que les podrán cuidar en el futuro? Un estudio basado en un modelo mixto de micro-simulación y basado en agentes (ABM)" (CSO2017-89721-R).

https://doi.org/10.1016/j.alcr.2022.100465
Received 14 January 2021; Received in revised form 18 January 2022; Accepted 4 February 2022
Available online 8 February 2022
1040-2608/© 2022 Elsevier Ltd. All rights reserved.

happens because exploitation of this detailed information is not always easy. But at a more fundamental level, and even in cases where the complete educational history is available, how to go about determining the level of education that conditions behaviour is not always obvious, especially if the date of completion of an individual's highest studies comes, for example, after he or she enters into a relationship or the birth of her or his children. Should the analyst consider this highest level of education or, instead, the level reached before the first stages of family formation?

Imprecision in the definition of educational level is problematic, especially when one analyses the relationship with lifecycle transitions which occur in the same age span as formal studies. This is the case, for example, with events like leaving the parental home, forming the first union, or birth of the first child, which often occur while studying, or are frequently postponed until after the end of studies.

For example, numerous works have shown that first childbearing is dependent on the duration of studies because fertility risks are very low for women while being enroled (Blossfeld & Huinink, 1991; Kreyenfeld & Konietzka, 2008). Following this line of thinking, some authors have even stated that the effects of education on family formation behaviour are often more a function of the age at the end of studies than the results of changes in preferences associated with reaching a higher educational level. For example, Ní Bhrolcháin and Beaujouan (2012), Neels, Murphy, Ní Bhrolcháin, and Beaujouan (2017), and Raymo (2003) argue that the increase in the duration of studies merely postpones behaviour like forming a first union or having a first child and therefore affects their tempo much more than their quantum, to use standard demographic terminology (Bongaarts & Feeney, 1998). Therefore, having information on the age at the end of studies may be equally, or even more important, than knowing the final level of studies. When the data source provides only the maximum level, as is frequently the case, the analyst may be forced to impute this age or duration using a standard educational age schedule. But individuals with the same degree do not necessarily have the same educational history (Kravdal, 2004), and often individuals do not complete their education at the standard age, mainly because of interruptions and repetitions (Baizan, Michielin, & Billari, 2001; Buchmann, 1989).

For example, in countries like Great Britain and the United States, young people often interrupt their studies to work for long periods, after which they return to school or university (Green & Green, 2008). In fact, around 80% of undergraduate students in the United States work while studying, most of them due to high university fees, but also to improve their employment conditions (Riggert, Boyle, Petrosko, Ash, & Rude-Parkins, 2006). This makes it difficult to determine which part of the life cycle is focused on studies and which on working (Kerckhoff, 2001).

Similarly, in Northern European countries there are provisions that compensate for deficiencies in the educational system and that give second or third chances to individuals who have dropped out prematurely (Andersson et al., 2009; Oláh & Bernhardt, 2008). A clear example is Sweden where the system is flexible with major public policies which ease re-entry and encourage people to resume their studies, even at late ages (Hoem, Neyer, & Andersson, 2006; Boucher, 1982).

By contrast, in Austria and Germany, interruption of studies and repeating subjects is not very common. This is partly due to the effort to gradually incorporate students into the formal labour market, ensuring that professional skills are also acquired during academic training (Muller, Steinmann, & Ell, 1998). The link between the educational system and the labour market is institutionalised through a "Dual Education System" in which students can combine studying and working for a few years (Petanovitsch & Schmid, 2016).

These disparities between countries are, to a certain extent, the product of national educational systems that tend to promote either homogeneity or diversity and competitiveness at the curricular level (Bulle, 2011) although, in some countries, social support is provided to continue education, help employability, and even to contribute towards the formation of a family (Breen & Buchmann, 2002).

In situations where studies are frequently interrupted, the use of the maximum level reached can lead to errors in measuring underlying risks, even if the age at completion is known rather than imputed. For example, the risk of childbearing while studying in the case of individuals who interrupted their studies could be overestimated, since they might have had their children during an interruption.

Another subtler problem is that if the age or year of the end of studies is not known, and if it is supposed that educational careers are continuous when in fact they are not, then the analyst risks altering the true order of events and, for example, might assume that an individual will have reached a specific educational level before the birth of the first child when, in fact, it is the other way around. Moreover, and in a more general way, taking into account only the maximum level is akin to considering it as a fixed characteristic, when education is an accumulative process and the maximum level reached changes with time as the person studies. This may lead to an erroneous anticipatory analysis or, in other words, explaining current behaviour in terms of an educational level that will, in fact, be reached in the future (Hoem & Kreyenfeld, 2006).

Another common shortcoming of the information about the educational level provided in numerous surveys or censuses is that the maximum level considered is the one successfully completed. This adds to the above-mentioned problem of ignoring periods of non-studies the real possibility of discarding the final part of the studies, if people do not end up with a diploma or a degree. Again, it can lead to biases in the measurement of relative risks of events related to family formation and enrolment status if individuals are classified in the group of people who have finished their studies at a certain age, when they were still studying at that age though did not obtain the qualification afterwards.

In this work, we measure some of the biases introduced by these kinds of limitations and try to address some of them with a proposal for a new questionnaire on education that could help to clarify some of the confusions listed above. In general terms, our proposal is closer to the view that it is more important to focus on the duration of studies than on the final educational level. However, using duration alone is not enough when exploring the relationships between education and other events of the lifecycle, because, for this, one needs to take into account discontinuities in the educational process and also, and more fundamentally, to consider their meaning in the life of individuals. This explains why we do not exactly use duration but, rather, age at the end of studies. In fact, we propose to define and use two different ages at the completion of studies in order to better explore the interactions of education with events related with the transition to adulthood. This, in turn, could make it possible to define duration of studies more thoroughly, taking into account its two different dimensions: as the main period of enrolment, or, more generally, the period of life devoted to studies, and absolute number of years spent studying, which is an alternative way of measuring the educational level.

The two indicators we define and compare in this work are *age at the end of the main studies* (AM) and *age at the end of the highest studies* (AH). The former builds on a proposal by Dourleijn, Liefbroer, and Beets (2002), while the latter is closer to the standard way of measuring the educational level. In the remainder of this work, we explain why differentiation should be made between these two ages and define the exact methodology for estimating them if the source used includes the complete educational history of individuals. In theory having this history should be the best situation for the analyst. However, as we will argue, the number of questions necessary to reconstruct it is very high, which leads to major complications in their collection and use.

For example, in the Fertility and Family Surveys (FFS) conducted in the 1990s, the last international survey of this kind which asked for the complete educational history, the number of questions asked is ten (Supplementary Table 1). Among them, were questions about the highest level of education satisfactorily completed, starting dates, the date on which each of the studies was completed or abandoned, whether

the studies were part-time, etc. Due to the complexity of the questionnaire, only 6 of the 23 countries implemented the survey completely; and 7 partially. However, they included the questions on the start and end dates of the study periods. This explains in part why recent fertility surveys, for example the Gender and Generation Surveys (GGS), generally opt for much simpler and less detailed questions for defining the educational level, which greatly simplifies their use and guarantees that the information obtained will be the same for all the countries concerned. Unfortunately, in the case of the GGS, the gain in simplicity is obtained at a great cost, because it no longer allows researchers to ascertain whether studies were interrupted or abandoned. In order to overcome these two opposite kinds of problems, excess complexity and excess simplicity, we propose a new set of questions intended to help in providing the maximum possible useful information on the level and duration of studies, taking into account the most important and relevant interruptions, with the minimum number of questions, which would be much less numerous and easier to use than when the survey collects the complete educational history.

2. Definition of two different ages at the end of studies. Relationship with the educational level as a factor influencing social behaviour

Generally speaking, education can affect behaviour like that associated with transition to adulthood through two dimensions. The first is related with changes in preferences brought about by the higher educational level. This is the viewpoint generally favoured by economists and sociologists (Barro & Lee, 1993; Becker, 1964; Boudon, 1974). The second is linked with the duration of studies, which can affect the life cycle of individuals by postponing the main stages of the transition to adulthood. This is the angle we favour in the present study, although the number of years studied can also be used as an alternative way of defining the educational level, as argued by Dourleijn et al. (2002).

In order to help measure the duration of studies from this perspective, we define *age at the end of main studies* (AM) which is different from the more commonly used *age at the end of highest studies* (AH). In technical terms, AM is defined as the age at which individuals end their last educational period with no interruption of more than 18 months, which means in practice no interruption of at least two academic years. The underlying/motivating hypothesis is that the individual's life cycle can be divided into at least two stages (Diagram 1): the first is focused on studies, and the second on the processes linked to family formation and entering the labour market. In this regard, we suppose that the first long interruption of studies, of at least two consecutive academic years can, for most individuals, represent the start of this second stage.

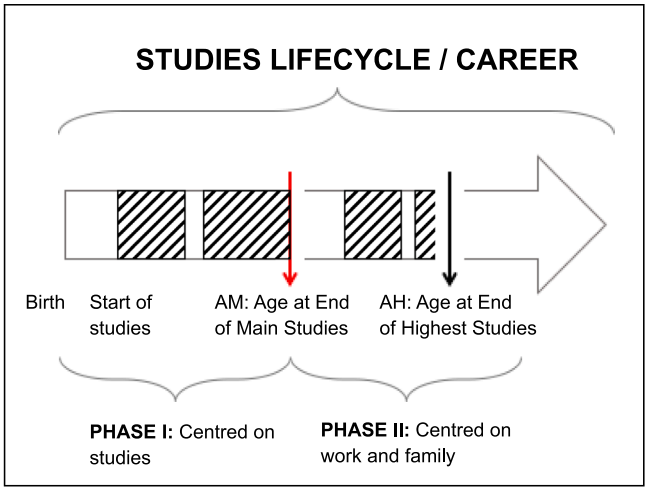


Diagram 1. Age at the end of main studies (AM) and at the end of highest studies (AH), in relation with lifecycle stages.

Determining an age that represents the ending point of the life stage centred on studies also allows us to obtain the educational level reached at that age, which is also prior to entering the labour force and to the start of the family formation stage. This is important for the study of the causal relationship between education and behaviour insofar as we are now considering the educational level that really affects future family formation events, and not just potentially, depending on whether the individual will resume his or her studies afterwards. We therefore avoid the danger of anticipatory analysis (Hoem & Kreyenfeld, 2006). In this sense AM should be the basis for the definition of the main educational level and not AH, because if the two are not equal, the latter will probably occur at the same time or even later than most of the events of interest potentially conditioned by the current educational level.

We also define and estimate age at the end of the highest studies (AH), which is associated with the highest educational level, and corresponds with the traditional way of measuring this level. However, we define it in a slightly more precise way, taking into account the end of the last period of studies, whether or not the person obtained the qualification, as we do also for AM. AH often coincides with AM but, as we will see below, when they are not equal, the difference between the two can be very high indeed. If this is the case, using AH to determine the duration of the period of studies, or the educational level, can lead to questionable results, for the reasons discussed above, and also because the real duration of studies will actually be shorter due to the interruptions.

For most individuals, there will be a direct correspondence between AM and the number of years studied, as this is the case when there are no interruptions or repetitions of a course. Accordingly, AM can be used directly to estimate the educational level associated with the main studies, which is not the case for AH, because of interruptions. This can be a boon when comparing between countries, especially when their educational systems differ, as was shown again by Dourleijn et al. (2002) when they surveyed the educational level of 17 European countries with FFS data. They compared the results of using duration of studies for deriving educational levels with levels obtained from the ISCED classification. For this purpose, they defined what they call a 'career--indicator', equal to the age at the end of the studies prior to an interruption longer than 18 months, and this is precisely AM. They showed that levels based on that age allowed them to obtain more meaningful results when comparing between countries. For example, they measured the degree of association of educational distribution between countries and observed that regional patterns were better recovered, and differences between countries were smaller with their 'career-indicator' than when using ISCED levels. They concluded that using ISCED levels led to "questionable" results. Festy and Prioux (2002) reached the same conclusion in their evaluation studies of the FFS surveys. The kind of problems they found is a consequence of differences between countries both in the educational system and the way in which the questionnaire was used to obtain the ISCED levels.

Taking interruptions into account is also crucial for measuring the risks of making certain lifecycle transitions while studying, by comparison with situations of non-studies. In order to demonstrate this, Kravdal (2004) and Zabel (2009) used surveys providing the educational history of individuals and applied them to measure these relative risks as accurately as possible. They then measured these relative risks from the same survey again, but as if the only information available was the highest educational level reached, and thus discarding the educational history.

For example, Kravdal used the Norwegian population register for modelling the risk of having children for women born in 1969, depending on whether they were enroled at each age and on their educational level at that age. He estimated the model using the complete educational history of each woman, thus taking into account all the interruptions there might have been. He then assumed that only the final educational level was available and used different imputation

procedures to determine, on the basis of their final educational level alone, whether each woman was enroled up to 28 years of age. He noted that the differences were very important: if the full history is used, the mean value of this relative risk is 32%, compared with women who do not study, but it varies between 28% and 54% when different methods of imputation are used.

Zabel did a similar exercise using data from a West German survey that also provides the complete educational history. Again, she compared the results of measuring the relative risk of fertility of women while studying based on this detailed data and different imputation procedures, assuming that the only information available would have been the final educational level. Unlike Kravdal, she took into account the first interruption of studies to define an indicator similar to our AM. And also, unlike Kravdal, she observed that the estimates of relative risk in the three different kinds of imputation she devised do not differ significantly from the values obtained with the full history. Zabel suggested that this is due to the fact that in Germany, interruptions, or repetitions of the same course, or drop-outs, are rare. There are, therefore, very few differences in the educational history of people with the same level of education, so having only the maximum level allows a correct reconstruction of the educational history in most cases.

These results indicate that having the complete educational history is not always necessary. In a country like Germany, educational paths are homogeneous, and interruptions are rare. By contrast, in Norway, there is a greater diversity of situations because it is common for students to interrupt their studies. In a later section we conduct an analysis similar to those of Kravdal and Zabel, with a broader scope, analysing three lifecycle events, comparing between twelve European countries and considering men and women separately.

3. Data and methods

This study has four main objectives. The first is to determine to what extent the duration of studies and our two ages at the end of studies can be measured from information collected in the main fertility and family surveys and, for this, we used the questionnaire and data from eleven of them. Three of these surveys are internationally comparative: the FFS, conducted in the first half of the 1990s in 21 European countries, as well as in the United States, Canada and New Zealand; the Demographic and Health Survey (DHS), which is the world's leading survey of this type, with over 90 countries and more than 300 surveys from the 1980s to the present; and the GGS, the successor to the FFS, with 19 countries and more than 40 surveys or waves, starting in year 2005 and continuing up to now. Five of the other surveys used are Spanish, from 1985 to 2007, and three surveys are Mexican, from 2003 to 2017. With this selection we seek to show different styles of questions about education in the data sources, to reflect the characteristics of each one of them and, with this, to analyse whether their content allows an estimate of the duration of studies and, if this is not feasible, what type of substitute information needs to be considered, and the extent to which our two ages of completion of studies can be obtained directly or whether they need to be imputed with the information available in the survey.

The second objective is to show that differentiating between "main" studies and "maximum achieved" studies is important in the context of analysing the relationship between education and family formation. Hence, we will see that if the survey data only allow for obtaining AH, this can lead to results with a considerable bias. In this case, we compare both ages using the FFS data, as this is the only internationally comparative survey that allows reconstruction of the complete educational history.

For the third aim of this study, we used our two ages at the completion of studies to show how they compare when measuring the propensity to experience lifecycle events associated with the transition to adulthood, depending on whether people are studying or not. Here, we use a statistical model of survival analysis. This allows us to measure the relative risks associated with each event, while studying or not, before and after AM. The events considered are leaving the parental home, first union, and first parenthood. We used FFS data again for the twelve European countries for which the educational history data were sufficiently complete for our purposes.

Finally, our fourth objective is to present a proposal for a questionnaire on education that could make it possible to obtain our two ages at the end of studies, as well as more specific information on success and on the field of studies, with a minimum of questions, and without having to collect the complete educational history. We take as a basis the best questions of the analysed questionnaires and the experience of other authors when identifying the most important variables in the absence of complete educational history and in such a way that the imputations that must be made are easy and give the best results.

4. Analysis of the information on education contained in the surveys selected

AM can easily be derived from surveys if they allow reconstruction of the complete educational history, as is the case for example with the FFS. However, it is more common to find data sources that collect only the last educational level attained at the time of the survey, or less frequently, both the level and the date of completion of such studies. In order to analyse them better, we classify the surveys on the basis of the following three criteria (Table 1): availability of partial or complete educational history or only the final level of studies; availability or not of the date of completion of studies (or age); and as regards the level of education, whether it is that attained or only that which is satisfactorily completed, i.e., with the final qualification for the cycle of studies undertaken. This makes it possible to create 5 groups of surveys, ordered from having the most to the least detailed information on educational trajectory (we are unlikely to find a survey that belongs to the sixth group that we can create from these three criteria).

In order to have a reference point when comparing surveys, we took the example of the study path of a typical person, even if not overly common, and indicated what information each survey would provide us with. As represented by the upper panel of Diagram 2, our ideal person satisfactorily completes middle or secondary education and obtains the corresponding certificates, whereupon she takes a break that lasts less than one year. She then resumes her studies and continues until reaching the second year of university studies when she again interrupts her studies, this time for more than 2 years, during which she possibly begins her working and family life. At the end of this period, she resumes

Table 1
Classification of various fertility and family surveys according to the information available on educational history.

Information about the highest level of studies	Date of periods of study		
	Complete educational history	Only final level of studies	
		With date of end of studies	No date of end of studies
Level reached with or without degree	**(1)** FFS, SP91 and MX17	**(2)** SP06	**(3)** DHS, MX03 y MX18
Only level reached with a degree		**(4)** SP99, CAT07 y GGS	**(5)** SP85

Note: FFS: Fertility and Family Surveys; SP91: Spanish "Encuesta Sociodemográfica de 1991"; MX17: Mexican "Encuesta Demografica Retrospectiva 2017"; SP06: Spanish "Encuesta de Fecundidad, Familia y Valores de 2006"; DHS: Demographic and Health Surveys; MX03: Mexican "Encuesta de Salud Reproductiva de 2003"; MX18: Mexican "Encuesta Nacional de la Dinámica Demográfica de 2018"; SP99: Spanish "Encuesta de Fecundidad de 1999"; CAT07: Catalan "Enquesta Demogràfica de Catalunya de 2007"; GGS: Generations and Gender Surveys; SP85: Spanish "Encuesta de Fecundidad de 1985". Note that FFS, GGS and DHS are the official acronyms for these surveys, while the rest, for the Catalan, Spanish and the Mexican ones, are our own creation, for identifying the country or region of origin.

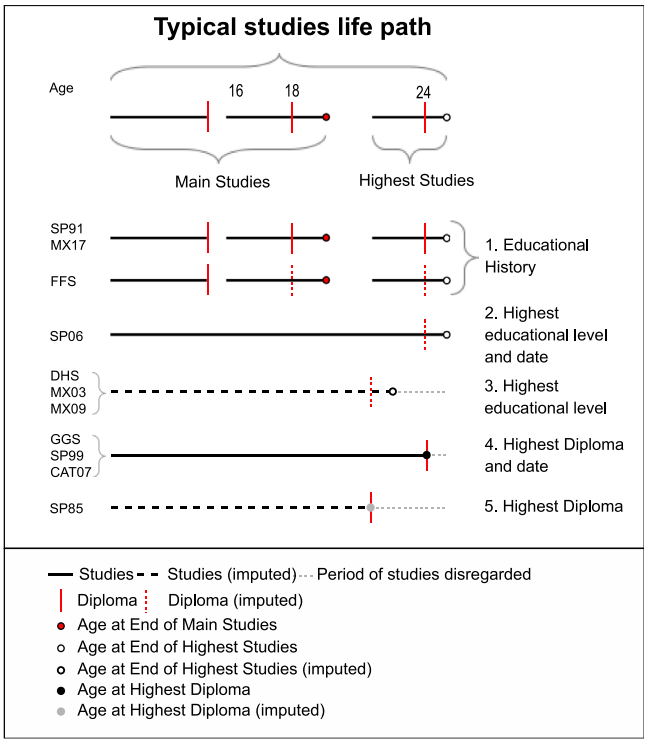


Diagram 2. Information obtained in 11 fertility and family surveys from a typical educational path.

her university studies, finishes her graduate studies, obtains a degree, and then begins a master's degree, but does not complete it. In the lower panel of this diagram, we represent the information obtained from the questions of the analysed surveys, with emphasis on whether our two ages of end of study are directly obtained, or whether their value must be imputed. The first group is made up of the three surveys that provide the complete educational history of individuals, the FFS, the SP91 and the MX17. In these surveys, people are asked about each of the episodes of study carried out, although with the FFS it is not possible to know the date of a diploma if it does not coincide with the end of an episode. In contrast, the SP91 and the MX17 offer the highest degree of detail, with the date of each diploma, the date of beginning and end of each episode, and whether it ended satisfactorily or not. With this type of survey, we know the duration of all the interruptions between the periods of study, and whether or not the person was studying when the events related to the transition to adult life occurred. Likewise, we can directly obtain our two ages, the main study completion age (AM) and the maximum age at study completion (AH). It is important to note that MX17 is a more complete survey than the others since it collects complete life histories of people, not only for education but also for family, union formation, fertility, working life, etc. However, and unfortunately, it is not possible to obtain AM with the surveys of the four other groups, either directly or even by imputation, since they do not include questions about possible study interruptions. Imputation of this age is impossible because, even if we could consider the cases of people who finish their studies at an age higher than the theoretical one, we could not know whether they either repeated a school or university year or interrupted their studies. In the second group, we have only the SP06, a survey that asks about the maximum level of education achieved and the year of completion of the studies. This data source has an advantage over the following groups (3, 4 and 5) in that it allows AH to be determined directly, although not AM. Third, we examined the DHS, the MX03 and the MX09 where only the maximum level and degree of the studies achieved are collected and therefore AH has to be imputed.

In the fourth group, the three surveys, SP99, CAT07 and GGS, ask for the highest completed educational degree and the corresponding age or date. The main problem with this type of question is that the person may have undertaken subsequent studies that he/she did not finish but this is not recorded in the survey and hence the value of AH obtained may underestimate the correct age in many cases.

In the fifth group, which corresponds to the scarcest possible information, we have the SP85, for which only information on the last completed educational degree was collected, without specifying the date when it was obtained. Again, the AH may be imputed but, as in the previous group, one cannot know whether the person undertook subsequent studies that he/she did not complete, leading in many cases to underestimation of its real value.

5. Comparison between the age of completion of the main studies and the maximum age

In Fig. 1, we compare our two ages of completion of studies based on FFS data for the twelve European countries which provided complete or partial information sufficient to reconstruct the educational history of individuals.[3] We aim to show that absence of data on the periods of interruption and the date of completion of studies can produce a distortion in the results if the highest age AH, and not the main age AM, is considered as an indicator of the level of education. The graphs for each country represent the distribution of the difference between the values of these two ages. We group the latter according to the value of this difference and of three educational levels based on the age AM of individuals when they completed their main studies: low level (if they completed them before the age of 17), medium level (between 17 and 20) and high level (over 20). We observe there are large differences between these ages and also between countries, which confirm the results of Kravdal (2004) and Zabel (2009). Thus, interruptions of several years of study were common in Scandinavian countries at that time, where this involved about 40% of students, especially Swedish women and Finnish and Norwegian men. In contrast, in Austria and Poland more than 90% of students finished their studies without interruption. There is no clear pattern of interruption according to the main educational level because, while in Austria, Estonia, Latvia and Lithuania individuals with a low educational level were the ones who most commonly resumed their studies, in the rest of the countries people with a medium educational level were likely to be the ones who interrupted them most frequently. In both cases, however, we observe that the process of resuming studies leads to a reduction in differences in educational level relative to the more highly educated, which a priori seems to be the opposite of what we would expect according to the so-called "Matthew effect" (Blossfeld & Maurice, 2019). But the observation that, in some countries, it is those with the lowest level of education who are most likely to return to formal education after an interruption does not take into account the fact that the improvement of vocational knowledge very often takes place outside the education system. In this sense (Blossfeld, Kilpi-Jakonen, Vono de Vilhena, & Buchholz, 2014; OECD, 2021), comparative studies of "adult learning" in Western societies, which provide a useful context for our analysis, have observed that people with university degrees are more likely than less educated people to subsequently improve their knowledge outside the formal educational system. This may, in fact, contribute to reinforcing knowledge inequalities throughout the professional life cycle,

[3] Six countries included the full set of 10 questions in their questionnaire, six more discarded some questions, on the subject of studies, whether the current period was part-time, and whether the person was studying or continued her or his studies after the age of 15 (supposed to correspond to the end of compulsory studies). Of the non-European countries, only the USA included enough information to reconstruct the educational history. Although we do not use this, it would be especially interesting to do so, as interrupted educational trajectories are very common in this country, as mentioned above.

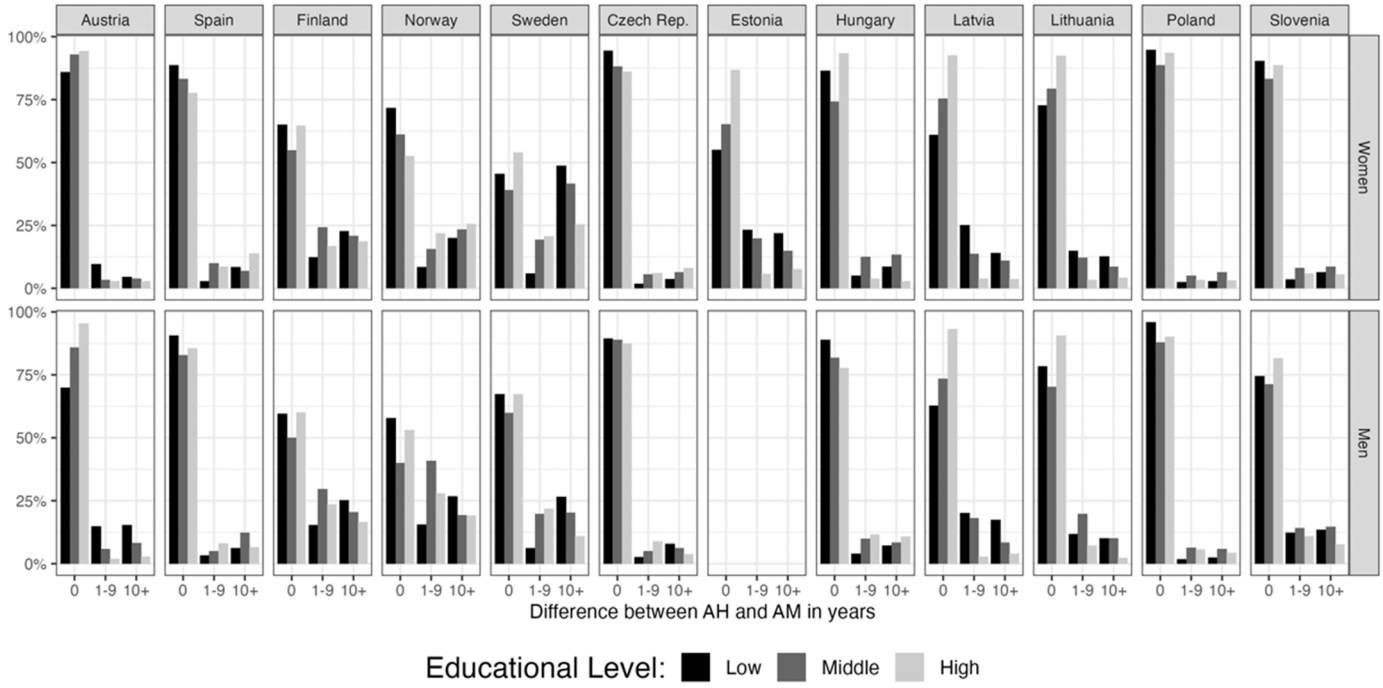


Fig. 1. Distribution of the difference between AM and AH for individuals aged 30–49 years, by sex and educational level, for 12 European countries, using FFS data, Note: The difference between AM (the Age at the end of Main studies) and AH (the Age at the end of Highest studies) is computed for each individual and distributed into three groups: *0*: same value; *1–9*: difference of less than 10 years; *10+*: difference of 10 years and more. For each group we separate individuals again into three groups: *Low*: people whose AM was less than 17 years; *Middle*: people with AM between 17 and 20 years; *High*: people with AM superior or equal to 20 years. The Estonian survey only includes data for women.

which would then correspond to a "Matthew effect".

In any case, the results show that interruptions were frequent, and that there were also great differences between countries. This means that in surveys such as GGS, with poor information on the educational process, it might be impossible to estimate correctly the level of relative risk of events related to transition to adult life for those who study, and the bias may be important when comparing between countries (as we shall show in the next section), as well as between educational levels.

6. The relative risk of events related to the transition to adulthood

In this section we present results from discrete-time survival models (Allison, 2014) in which a person's life cycle is broken down into short, quarterly episodes (we discarded use of the monthly period, because of its excessive computational cost, without any improvement in the results). For each episode we determined whether the person was in the main study phase or a subsequent phase, studying or not, and whether certain lifecycle events had occurred in the episode under consideration. Specifically, we measure the relative risk of three events associated with transition to adult life: leaving the parental home, first union, and first parenthood (backdated nine months to refer to the time of conception, in order to avoid reverse causality in cases of people who interrupt their studies due to the start of a pregnancy while enroled in education), comparing people who are studying with those who have finished their studies (to be precise the estimated coefficients are odds ratios but as the incidence risks, calculated for quarters, have values below 2%, the odds ratios and the relative risks can be considered as equivalent, since they will differ by less than 1%: see for example Davies, Crombie, & Tavakoli, 1998).

As indicated in Diagram 3, we considered two separate models. In the first we calculated the relative risk of experiencing one of these three events by assuming that individuals are studying continuously up to the maximum age AH. This type of model corresponds with the traditional

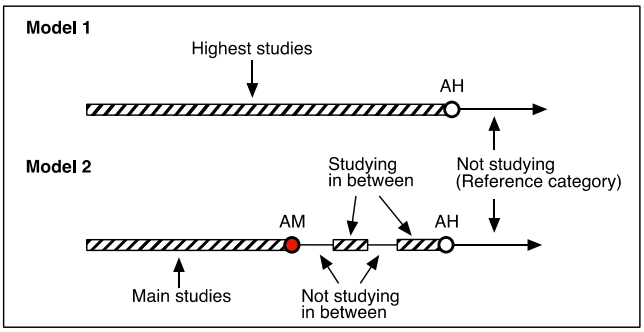


Diagram 3. Two ways of modelling the study period according to whether or not interruptions are considered.

analysis, erroneous in our opinion, in which interruptions are not considered and it is implicitly assumed that the educational level that conditions people's life is the maximum level reached. In the second model, we analyse in a more detailed way the individual educational path, which allows us to calculate the relative risks corresponding to actual periods of studies. Specifically, we compute the risk for the period of main studies which ends with age AM. We also break down the period between AM and AH, between periods of study and non-study, and calculate the separate risks for both types. In these two models, the reference category is the life-cycle period after the maximum age AH, and we eliminate the effects of the age pattern by introducing 'age' as a categorical control variable in these models.

With the results reported in Fig. 2, it can be seen that there were indeed significant differences in relative risks while studying for these three events, depending on whether we consider the period that ends with the main studies (variable *stud_main*) or the longer period that ends with the maximum or highest studies (variable *stud_max*). The largest differences are observed for women and for the first birth, especially for

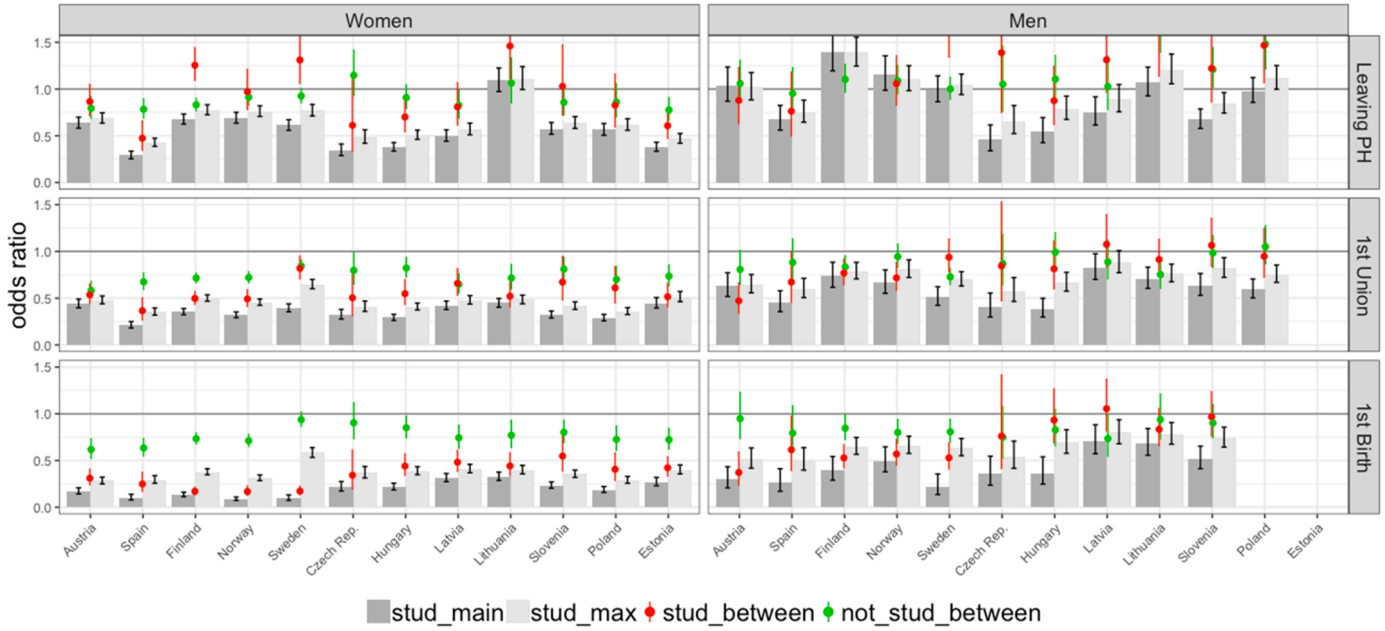


Fig. 2. Relative risks of leaving the parental home, forming a first union, or having a first child while studying, for 12 European countries, using FFS data, Note: The whole survey sample is used for each country, which means, in most cases, persons in the age range of 15–49 years. First births are backdated to the time of conception, i.e., nine months before. We compare here the results of the two models described in Diagram 3. In the two models the baseline risk for each type of event corresponds to the same period of life: after completion of all studies, i.e. after age AH; *stud_max* is the relative risk of occurrence of the three events analysed during the period of studies which ends with AH, as estimated with Model 1; comparable results for Model 2 corresponds to the combination of the three following parameters: *stud_main* is the relative risk that each event occur in the period before age AM; *stud_between* is the relative risk of occurrence in the periods of study between ages AM and AH, for people who interrupted their studies for more than one year; *not_stud_between* is the relative risk for the same individuals during the non-study periods between ages AM and AH. For each risk 95% confidence intervals are shown. Observe that the Estonian survey did not include data for men, and for Poland the number of births reported by men was too small to be usable.

Scandinavian women. This allows us to demonstrate again, as Kravdal did for first parenthood in Norway, that if interruptions are not considered, there is a clear danger of obtaining biased values for the risks associated with the three transitions, during the study phase. Another important result from the standpoint of this paper is that not all study phases have equal effects on the transitions analysed. Hence, the relative risks during the main study phase are clearly lower than during the study phases following the first long interruption, and this is so for all the events considered (comparing the level of the variable *stud_main* with *stud_between*). The only exception is for Austrian men who showed a higher risk of leaving the parental home and of first union during the main study phase than in subsequent ones. This result is a confirmation of the validity of the approach presented in this work, in which we distinguish between a life-cycle phase focused on studies, and another subsequent phase, after the first long interruption, in which studies and the beginnings of family formation processes are intermingled.

If we look more closely at the results of this figure, we can observe that:

- Being enroled inhibited fertility more than the formation of a union and even more than leaving the parental home. Likewise, studying was a more powerful inhibitor for women than for men.
- In general, those who studied were less likely to make the three transitions analysed, except for men from Scandinavian countries, Lithuania and Poland, as well as women from Lithuania, who were more likely to leave the parental home while being enroled, especially in the phase following the first long interruption.
- The effect of enrolment was very high for first motherhood, with very low values for the relative risks for women both in the main study phase and after a long interruption. In contrast, this inhibitory effect was much lower for men.

7. A proposal for a questionnaire on education in fertility and family surveys

As shown above, it is of great interest to take interruptions into account if we want to measure the real effect of studies on family formation processes, but also, and even more importantly, if we want to define the educational level that really conditions demographic and social behaviour at the start of adult and family life. Accordingly, it would seem to be beneficial to collect the complete educational history, as was done for the FFS. But as noted above, its implementation was a partial failure, as only 6 countries used the full version of the questionnaire and 7 used it partially in such a way as to provide sufficient information to reconstruct the educational lifepath. This heterogeneity and the sheer complexity for the analyst of dealing with the full questionnaire, explain why use of this information was very limited, especially for comparative studies. It would then be interesting to have a questionnaire with a minimum of questions, easier to use and which, on the other hand, allows the researcher to obtain the maximum amount of information by approximating what would be obtained from reconstruction of the complete educational history.

According to the argument of one the two authors mentioned above (Zabel, 2009), if it is not possible to have a questionnaire that allows reconstruction of the complete educational history, it is recommended that two types of information should be asked for: age or date when the highest studies were completed, and the highest educational level achieved. But this recommendation is based on the specificities of Germany's educational system. Kravdal (2004), on the contrary, calls for the inclusion of much more detailed information on the educational process.

In the present study, two aspects are considered to be essential, and yet are rarely collected: age or year of completion of studies without a long interruption, and whether or not these studies were completed with a qualification. In accordance with the previous results, one can observe

that there is frequently a difference of several years between the age at which a person completes his or her highest-level studies and the age at which he or she interrupts them for a long period of time, so it is essential to include questions that allow detection of these prolonged interruptions. In addition, knowing whether studies were abandoned without a qualification is important, since numerous authors have shown that dropping out may affect the subsequent lifecycle of individuals (See for example Ripamonti, 2017 for a recent review of the literature).

Another type of information that is important, is the highest level of education achieved. This level has been and is used in a large number of surveys as the only information on education. But since each country has its own education classification, it is usually recommended that researchers should use a classification that is simple and that facilitates comparison between countries, for example the ISCED ranking. But the experience is that despite the fact that this classification was created for comparative purposes, the results are more inconsistent than those obtained with a classification based on length of study, as discussed above.

After reviewing the questionnaires of the surveys mentioned above and, on the basis of our results as well as findings of other authors, we have identified the minimum information that each fertility and family survey should have in order to improve the quality of the analysis. We have found in previous sections that having information on interruptions, age before the first long interruption, and educational level reached is of the utmost importance. However, the needs of each country are different and resources are always limited, so having a questionnaire as broad as the one proposed by the FFS is debatable not only for reasons of budget or balance with the other sections of the survey, but also because of ease of use, since managing a complete educational history is complicated. We therefore propose a module of 12 questions (Supplementary Table 2) which allow direct access to main age and maximum age. In cases where there are restrictions of space or budgetary resources, an alternative proposal has also been designed with only 6 questions (Supplementary Table 3). In both proposals, a distinction is made between main studies and final or maximum studies achieved, and the age of completion of both is obtained. In turn, in the full questionnaire, reference is made to the type or field of studies, both those interrupted for a period of at least one academic year and those of a higher level. The main advantage of the complete questionnaire is that the exact date of achieving qualifications is obtained even in cases where it does not coincide with the date of completion of studies. With the reduced questionnaire, the date must be imputed. Diagram 4 specifies, like Diagram 2 for existing surveys, what kind of information could be obtained from our two proposed questionnaires.

For those individuals who have not interrupted their studies, the full questionnaire would be completed by question 7, while for those who have suspended their studies for a period of more than one academic year, the questionnaire continues in question 8 (date of the interruption) and, in the case of having obtained an academic degree prior to this period of time, only two additional questions would have to be answered, these being related to the level obtained prior to this episode. In cases of not having obtained an academic degree before the interruption, the questionnaire collects information regarding the educational level that was not completed, and the date of the last degree satisfactorily obtained.

If there are resource constraints, the reduced proposal is the more appropriate (Supplementary Table 3). With it, as with the full proposal, it is possible to differentiate between the main studies and the final or maximum studies achieved. In addition, it allows collection of the three types of information that have been identified as fundamental throughout this work. First, the educational level reached; second, the date on which these studies were completed; and, third, information referring to interruptions greater than one academic year. This is possible because the date and the last level being studied are collected, which also makes it easier to impute the date of the last academic degree obtained.

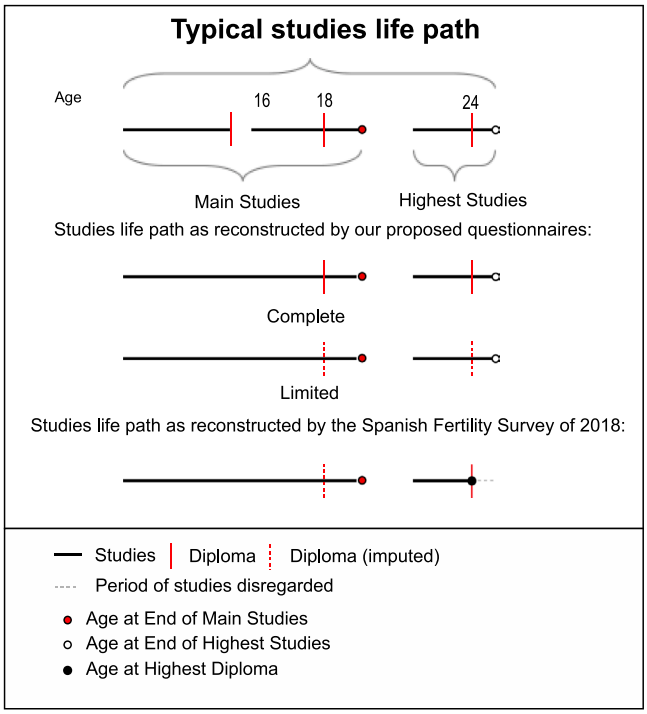


Diagram 4. Information obtained from the two versions of the proposed questionnaire and from the questionnaire of the Spanish Fertility Survey of 2018.

In fact, as shown in Diagram 4, the main difference in the information collected is that with the complete questionnaire, the date of obtaining the qualification is available in all cases, whereas with the reduced questionnaire, this date has to be imputed. In the latter situation, the privileged information is the number of years studied when, with the complete questionnaire, information is also obtained on qualifications, as well as courses that were not successfully completed. Moreover, the fact that no questions are asked about obtaining a qualification in the short questionnaire makes it impossible to ascertain directly whether the person has abandoned his or her studies, especially if he or she did so in the final year of an educational cycle. As mentioned above, the short questionnaire neither asks about the field of studies, a factor that has been shown to be related to later family behaviour (Hoem et al., 2006).

Finally, the main limitation of these two questionnaires is the assumption that there is only one long interruption, of more than one year, after completion of the main studies. The FFS data again allow us to check the degree of such an approximation. Thus, for nine of the twelve countries analysed in Fig. 1, more than 90% of the people who interrupted their studies for more than one academic year did so only once. This simplification is more debatable only for the three Scandinavian countries analysed (Finland, Norway and Sweden) where around 20% of people with long interruptions had at least two. Another limitation of these questionnaires is that they do not allow direct determination of the duration of this long interruption, so this has to be imputed from information on the final studies.

This diagram also presents the educational lifepath, as collected by the 2018 Fertility Survey of the Spanish Statistical Institute (INE). Its education module is a hybrid between the 1999 Survey questionnaire (SP99 in Table 1) and the reduced version of our proposals (Supplementary Table 4). The INE had an objective of comparability with its previous survey, but was open to the needs of researchers, which allowed for the introduction of questions necessary to obtain our age AM, although not of AH, but rather of the age at the last diploma, in accordance with the questions of 1999. However, it should be noted that

the questionnaire of 2018 is more thorough regarding the first long interruption, by asking about the reasons for and the duration of it. It also includes a question about the field of study, at the request of other researchers.

8. Discussion

The results of this research, in line with the work of Dourleijn et al. (2002), Kravdal (2004) and Zabel (2009), show that the educational path of individuals is not necessarily uniform and once they have left school there is a possibility of resuming their studies, which is customary, for example, in the Scandinavian countries, and this also occurs more often among people with a low or medium level of education at the time of interruption. Hence, having only the last level of education attained at the time of the survey is insufficient for the study of transitions to adulthood.

Once the hypothesis that individuals interrupt their studies and later resume them has been tested, the information that seems to be important is the duration of studies going from childhood to the first interruption of these studies which, in this research is called age at the end of main studies. With regard to our recommended variable, we found three advantages over the use of this age compared with the maximum educational level achieved. (1) Individuals who have not yet completed an educational level will receive a score that reflects the fact that they have accumulated a number of years of human capital during their studies, even those who have not completed that level. For example, if an individual has been enroled in a university for a couple of years, the number of years of human capital is greater than for another individual who has not completed any years. The level of education attained does not take into account these subtleties by collecting only the last level of education completed, while age does. (2) The age of completion of the main studies can be more easily transformed into a variable that changes over time compared with the maximum level of education attained, because the latter does not take into account the possibility of interruptions. (3) This indicator allows for an equivalent measurement of the educational attainment of individuals in all countries studied, despite the great diversity of their education systems which rules out use of the final level of education. For example, in the GGS questionnaires all participants are asked about the highest level of education they have completed and are then classified according to the ISCED level, which does not reflect the accumulated human capital of those levels initiated but not completed.

On the other hand, we cannot deny that, theoretically, the most recommendable situation would be to have the complete educational history of the individuals. But this can have a very high cost in the implementation phase of the survey and in use of the information obtained, as demonstrated with the FFS. It should be noted that these surveys were carried out in the 1990s and, from the perspective of the usefulness of the proposed questionnaire, it would be interesting to know whether situations of interruption of studies are more or less frequent nowadays. The 2018 Spanish fertility survey, which uses a similar questionnaire, provides an answer. It allows calculation of the proportion of women interrupting their studies for more than two years. Results show that this proportion remained stable, at around 10%, for the cohorts born from 1960 to 1980 (INE, 2018), which is comparable to the level observed for Spain with the FFS.

In this research, the elementary variables for calculating the duration of studies have been identified, and in the case in which, for reasons of space or budget, it would not be possible to collect the complete history, we propose a simplified questionnaire that makes it possible to obtain the date of or age at completion of the main and maximum studies or, in other words, it takes into account the first main interruption. Such information would also considerably facilitate making imputations to recompose the educational trajectory of individuals and would make it possible to know the exact moment of the first long interruption of studies, as well as to ascertain whether it was the result of dropping out.

However, the substantive effect of education as a determinant of the stage of family formation does not seem to be given due weight in international comparative data sources. For example, the countries covered by the DHS, which has a very poor educational questionnaire, generally do not have national-level surveys containing the educational history of their population, even though the vast majority of the countries in which this survey is carried out have not yet completed their first demographic transition. Similarly, countries that are in the more advanced stages of their second demographic transition do not always have detailed surveys of educational levels and, unfortunately, the GGS presents very limited information on educational history, referring to the highest level of education completed without, for example, covering information on levels of education started but not completed, or the possibility of long interruptions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.alcr.2022.100465.

References

Allison, P. D. (2014). *Event History and Survival Analysis* (second ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Andersson, G., Rønsen, M., Knudsen, L. B., Lappegård, T., Neyer, G., Skrede, K., … Vikat, A. (2009). Cohort fertility patterns in the Nordic countries. *Demographic Research, 20*(14), 313–352.

Baizan, P., Michielin, F., & Billari, F. C. (2001). Political economy and life course patterns: the heterogeneity of occupational, family and household trajectories of young Spaniards. *Demographic Research, 6*(8), 191–240.

Barro, R. J., & Lee, J. W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics, 32*, 363–394.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press for NBER.

Blossfeld, H. P., & Huinink, J. (1991). Human capital investments or norms of role transition: how women's schooling and career affect the process of family formation. *American Journal of Sociology, 97*(1), 143–168.

Blossfeld, H. P., Kilpi-Jakonen, E., Vono de Vilhena, D., & Buchholz, S. (2014). *Adult Learning in Modern Societies: An International Comparison from a Life-Course Perspective*. Cheltenham, UK: Edward Elgar.

Blossfeld, H. P., & Maurice, J. (2019). Education as a Lifelong Process. In H. P. Blossfeld, & H. G. Roßback (Eds.), *Education as a Lifelong Process. The German National Educational Panel Study (NEPS)* (pp. 17–34). Berlin, Heidelberg: Springer.

Bongaarts, J., & Feeney, G. (1998). On the quantum and tempo of fertility. *Population and Development Review, 24*(2), 271–291.

Boucher, L. (1982). *Tradition and Change in Swedish Education*. Oxford: Pergamon Press.

Boudon, R. (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.

Breen, R., & Buchmann, M. (2002). Institutional variation and the position of young people: a comparative perspective. *Annals of the American Academy of Political and Social Science, 580*, 288–305.

Buchmann, M. (1989). *The Script of Life in Modern Society*. Chicago: University of Chicago Press.

Bulle, N. (2011). Comparing OECD Educational Models through the Prism of PISA. *Comparative Education, 47*(4), 503–521.

Davies, H. T. O., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *British Medical Journal, 1998*(316), 989–991.

Dourleijn, E., Liefbroer, A. C., & Beets, G. (2002). Comparing the 1988 International Standard Classification of Education (ISCED) with retrospective information from educational histories. In E. Klijzing, & M. Corijn (Eds.), *Vol. 2. Dynamics of Fertility and Partnership in Europe: Insights and Lessons from Comparative Research*. Geneva: United Nations.

Festy, P., & Prioux, F. (2002). *An Evaluation of the Fertility and Family Surveys Project*. Geneva: United Nations.

Green, A.G. & Green, D. (2008). Canada's Wage Structure in the First Half of the Twentieth Century (with Comparisons to the United States and Great Britain). Available at: ⟨https://economics.ubc.ca/files/2013/05/pdf_paper_david-green-canada-wage-structure.pdf⟩.

Hoem, J., & Kreyenfeld, M. (2006b). Anticipatory analysis and its alternatives in life course research. Part 1: the role of education in the study of first childbearing. *Demographic Research, 15*(16), 461–484.

Hoem, J., Neyer, G., & Andersson, G. (2006a). Education and childlessness: the relationship between educational field, educational level, and childlessness among Swedish women born in 1955-59. *Demographic Research, 14*(15), 331–380.

INE. (2018). *Encuesta de Fecundidad de 2018*. Madrid, Spain: Instituto Nacional de Estadistica.

Kerckhoff, A. C. (2001). Education and social stratification processes in comparative perspective. *Sociology of Education, 74*, 3–18.

Kravdal, O. (2004). An illustration of the problems caused by incomplete education histories in fertility analyses. *Demographic Research, Special Collection, 3*(6), 135–154.

Kreyenfeld, M., & Konietzka, D. (2008). Education and fertility in Germany. In I. Hamm, H. Seitz, & M. Werding (Eds.), *Demographic Change in Germany*. Berlin, Heidelberg: Springer.

Muller, W., Steinmann, S., & Ell, R. (1998). Education and labour market entry in Germany. In Y. Shavit, & Muller (Eds.), *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*. W. Oxford: Clarendon Press.

Neels, K., Murphy, M., Ní Bhrolcháin, M., & Beaujouan, É. (2017). Rising educational participation and the trend to later childbearing. *Population and Development Review, 43*(4), 667–693.

Ní Bhrolcháin, M., & Beaujouan, É. (2012). Fertility postponement is largely due to rising educational enrolment. *Population Studies, 66*(3), 311–327.

OECD. (2021). *OECD Skills Outlook 2021. Learning for Life*. Paris: OECD Publishing.

Oláh, L. S., & Bernhardt, E. M. (2008). Sweden: combining childbearing and gender equality. *Demographic Research, 19*(28), 1105–1144.

Petanovitsch, A. & Schmid, K.(2016). Dual vocational education and training in Austria, Germany, Liechtenstein and Switzerland: comparative expert study. Zurich: Secretariat, Donor Committee for Dual Vocational Education and Training.

Raymo, J. M. (2003). Educational attainment and the transition to first marriage among Japanese women. *Demography, 40*(1), 83–103.

Riggert, S., Boyle, M., Petrosko, J., Ash, D., & Rude-Parkins, C. (2006). Student Employment and Higher Education: Empiricism and Contradiction.

Ripamonti, E. (2017). Risk factors for dropping out of high school: a review of contemporary, international empirical research. *Adolescent Research Review, 3*(3), 321–338.

Zabel, C. (2009). Do imputed educational histories provide satisfactory results in fertility analysis in the west German context? *Demographic Research, 21*(6), 135–176.