Dropout and Degree Completion in Doctoral Study:

A Competing Risks Survival Analysis

Margaux van der Haert
Université libre de Bruxelles, SBS-EM, Brussels, Belgium

Elena Arias
Inter-American Development Bank

Philippe Emplit
Université libre de Bruxelles, OPERA - Photonique, Brussels, Belgium

Véronique Halloin
FRS-FNRS, Brussels, Belgium

Catherine Dehon
Université libre de Bruxelles, ECARES, SBS-EM, Brussels, Belgium
CP 114/04, Av. F.D. Roosevelt, 50. B-1050 Brussels, Belgium.

Abstract

In this article, the determinants of "time to dropout" from doctoral studies and "time to Ph.D. completion" are studied for a sample of doctoral candidates from the *Université libre de Bruxelles (ULB)*. We show that a discrete-time competing risks survival analysis is necessary to study this type of data. Results indicate that neither gender nor nationality significantly influences the doctoral path. Students with grants (FNRS doctoral fellowships, FRIA doctoral fellowships and ULB grants) have higher Ph.D. completion hazards than teaching assistants. Grants influence degree attainment even when controlling for the ability of the student. Students with FNRS doctoral fellowships and FRIA doctoral fellowships have lower dropout hazards than all the other students, although their ability plays a role therein. Finally, our findings suggest that there are no significant differences between fields of study for most types of financing, but that they exist for unfinanced students.

1. Introduction

Individuals holding doctoral degrees are of utmost importance to our current society. On one hand, the advent of the knowledge economy implies a strong need for researchers. In most countries, possessing a degree of Doctor of Philosophy ("Ph.D." or "doctorate") is a prerequisite for pursuing a permanent career as professor or researcher in institutes of higher education[1]. Moreover, doctorate holders also contribute to the progress of knowledge by working as researchers in the business or the government sector. On the other hand, doctorate holders are also valuable in jobs outside research. Indeed, thanks to their training in research, Ph.D. holders often have developed skills, such as managing and synthesizing knowledge, precise communication, and overcoming difficulties. All these skills and competencies are great assets to hold in high-level business positions. For example, the survey "Careers of Doctorate Holders" (CDH) in Belgium indicates that about 50% of Ph.D. holders do not have a job as researcher (Moortgat, 2011). At the same time, the dropout rate of Ph.D. candidates from doctoral education is high. For example, of those who started a doctorate in 2001 at the *Université libre de Bruxelles* (ULB), one of the largest universities of the French Community of Belgium (FCB)[2], 47% had dropped out of the program by 2010. This figure is comparable to the attrition rates reported in American studies (45% in Ferrer de Valero (2001), around 50% in Nelson and Lovitts (2001) and 40-50% in Smallwood (2004)).

In Europe, the support of research and doctoral education has evolved positively in the last decade but there are indications that it can still be improved. Very few countries currently attain the Lisbon Agenda goal of an investment of 3% of GDP in Research & Development, which was renewed in the Europe 2020 strategy. However, the situation is not good in Belgium where in 2008 this figure[3] was of only 1.92%, which is in fact close to the European mean. Although in Wallonia, the southern and French-speaking part of Belgium, investment in research has increased in the last years, particularly with the implementation of the "Marshall Plan"[4], the money goes mostly to applied research, leaving fundamental research underfinanced. However, this second type of research is the backbone of a knowledge economy and is necessary to ensure the continuous development of applicable innovations. Given the limited amount of financing on one hand, and the high dropout rate on the other hand, it is crucial to have an in-depth understanding of the doctoral process in order to improve the efficiency of doctoral study and allocate resources accordingly.

[1] In the International Standard Classification of Education (ISCED) maintained by the UNESCO, the Ph.D. level is one of the Advanced Research Qualifications of level ISCED 6. In the Bologna Process for the harmonization of higher education in Europe, a Ph.D. degree corresponds to a third cycle qualification.

[2] Belgium is a federal state with three regions (the Flemish Region, the Walloon Region and the Brussels-Capital Region) and three language communities (the Flemish Community, the French Community and the German-speaking Community). The communities are competent for the educational system.

[3] Figure from Eurostat (2011).

[4] The "Marshall Plan" is a publicly financed program for Wallonia who's objective was to invest 155 million € to enhance other things, the region's research and development related to the private sector (www.wallonie.be)

An efficient doctoral process is thus important to the actors who finance doctoral studies, such as universities, governments, firms and the FRS-FNRS[5] ("Fund for Scientific Research", or "FNRS" for its acronym in French). Besides financing individuals during their studies through assistantships and grants, universities invest in other types of resources such as time and cost of teaching during the doctoral curriculum (Bair and Haworth, 2005; Booth and Satchell, 1995; Siegfried and Stock, 2001; Smallwood, 2004). However, an efficient doctoral process is also important to students. Indeed, doctoral education is costly to them, in terms of time and money invested, foregone earnings and in some cases emotional costs (Nelson and Lovitts, 2001; Siegfried and Stock, 2001; Smallwood, 2004; Stricker, 1994). A reduction of the attrition rate and time to degree thanks to a better understanding of what determines failure or success at the doctoral level will encourage more people to consider the pursuit of a Ph.D. This is beneficial given the high value of doctorate holders in building a knowledge base economy.

This article aims to contribute to the understanding of the doctoral process in the French Community of Belgium (FCB). To achieve this objective, we analyze the determinants of "time to completion" of the Ph.D. and "time to dropout" of the doctoral process for Ph.D. candidates at the ULB using a discrete-time competing risks survival method. There has been a recent increase in research on the determinants of Ph.D. completion, attrition and time to degree of doctoral candidates, while this research had been scarce in the past (Abedi and Benkin, 1987; Gillingham *et al.*, 1991; Stricker, 1994). However, most of the literature on the determinants of the doctoral path uses data from the U.S.A., and to a lesser extent from the U.K. Notable exceptions are the studies of Seagram *et al.* (1998, Canada), van Ours and Ridder (2003, The Netherlands) and Visser *et al.* (2007, Flanders). To date, no quantitative research that we are aware of has been undertaken concerning the determinants of time to dropout and to Ph.D. attainment in universities of the FCB.

In addition to contributing to the understanding of the doctoral process, this article is also useful from an econometric point of view. Indeed, few authors who analyze data on education, and in particular on time to the attainment of the doctoral degree, use a discrete-time competing risks survival analysis. However, we will explain that this method is often needed to analyze the doctoral process quantitatively. In this article, the discrete-time competing risks hazard model of Scott and Kennedy (2005) will be used and presented in detail. It involves estimating, for each year after the start of doctoral study, the risk of dropping out of doctoral study and of completing the Ph.D. for a particular individual. The model will be applied to a sample of students who registered for doctoral study during the period 2001-2002 until 2009-2010 at the ULB.

[5] The FRS-FNRS finances fundamental research in universities of the FCB.

Time to withdrawal or Ph.D. completion can be determined by several factors. This study will focus in particular on the following aspects. First, we will investigate whether women are as likely as men to complete their doctorate in a timely manner. The same question will be studied for foreign students compared to Belgian. Second, the influence of the type of financing that a doctoral candidate receives on his probabilities of dropout and Ph.D. completion will be analyzed. It would be expected that students who receive no financing drop out more and graduate less often than their financed colleagues. Moreover, it is interesting to analyze whether there are differences between the types of financing, which differ for example in the length of time they are awarded and the duties they entail. Are students with teaching duties less likely to complete their doctorate and more likely to drop out? Do students with shorter financings complete their doctorate faster? In the context of limited resources for research, knowing which type of support is most effective is of high importance. Moreover, the link between type of financing and "ability" of the student will be studied. Indeed, as the types of support differ in terms of selectivity, one can wonder whether financing is well allocated, *i.e.* whether the students who receive the most selective financial aid are intrinsically the ones who will drop out less and complete their Ph.D. most often. In other words, we will investigate whether a particular type of support has high or low dropout and completion rates because of the ability of the students who receive it, or because of the characteristics of the type of financing received. Finally, the differences between fields of study are of interest too. The common belief is that students in human sciences are less likely to complete their doctorate than candidates in science and technology or health sciences, and have a higher time to degree if they do so. Moreover, these students have less easily access to financing. Indeed, during the period 2001-2009 almost 60% of candidates in human sciences at the ULB were mostly unfinanced during their doctoral education, while this percentage was of slightly more than 40% for health science students and 30% for science and technology students. This is partly due to the fact that one of the grants of the FNRS (the "FRIA" fellowship) is only accessible to doctoral candidates in these last two fields of study. Investigating whether there are really differences between areas of research is thus of particular interest.

This article is set out as follows. Section 2 presents the methodology and a review of previous research, Section 3 the system of doctoral education of the French Community of Belgium and Section 4 the data used in this article. Section 5 contains an exploratory analysis of the data and Section 6 a model-based analysis. Finally, Section 7 concludes.

2. Methodology and review of previous research

In this article, the methodology known as survival analysis is used to model, for each year, the hazard ("risk") of occurrence of two events, *i.e.* dropout of doctoral study and Ph.D. completion, for a particular individual. The question is thus *when* rather than *whether* these events occur. This is not the case for all articles that will be reviewed later in this section. Indeed, some researchers dichotomize their sample at some point in time to analyze *whether* an individual has experienced some event by that point or not. Mooney (1968) studies what determines whether a student attended graduate school during six to eight years without acquiring a Ph.D. for a cohort of Woodrow Wilson Fellows. Luneburg and Luneburg (1973) study the impact of fifteen predictors on thirteen criteria of graduate school performance, such as withdrawal and completion of Ph.D. in a four-year time frame, for the psychology doctoral program of the University of Washington. Ott *et al.* (1984) apply a legit model to predict retention in doctoral study at the start of the sixth year at the University of Maryland. Park (2005) studies associations between non-completion in a ten-year time frame and six factors in an English university. Stock *et al.* (2006) estimate first-, second- and two-year attrition with a probit model applied to data on 27 U.S. economics Ph.D. programs. Finally, Grove *et al.* (2007) analyze the determinants of success at three distinct stages of the economics Ph.D. program of Syracuse University. However, Willett and Singer (1991) explain why the "when" question is more interesting than the "whether" question. They argue that dichotomization can obscure knowledge about educational transitions, as it does not allow modeling the variation in risk over time. Moreover, contradictory conclusions can result from taking into account different time frames. They show that *"by asking when events occur, a researcher learns not only whether these events occur by each of several points in time but much much more"* (Willett and Singer, 1991, p. 408).

When the "when" question is analyzed, a survival model is the most appropriate method to use. Indeed, some students, called censored individuals, will not have dropped out or completed their degree when the period of data collection ends. A survival analysis deals with this problem of censoring[6]. Moreover, as Allison (1982) argues, it allows incorporating time-varying variables[7]. After the influential paper of Cox (1972), in which he developed the survival method to model human lifetimes, it gained in popularity in diverse disciplines, such as economics, sociology or engineering. However, most methods were developed for continuous-time data, the

[6] It is almost impossible to have a dataset on event occurrence without censored individuals. Singer and Willett (1993) discuss the different solutions to the censoring problem used by other authors, arguing that none of them deals with it as well as survival analysis.

[7] For a discussion of why standard multiple regression techniques cannot accommodate time-varying explanatory variables, see Allison (1982) and Flinn and Heckman (1982).

most popular one being the continuous-time proportional hazards model of Cox. In the perspective of adapting the survival method to educational data, Singer and Willet (1993) developed a discrete-time hazard model. Indeed, in an educational context like the doctoral process, the precise timing of an event is often unknown or time measured discretely. The model of Singer and Willett is based on the assumption that events can only occur at discrete times. An alternative is to assume that there is an underlying continuous-time model, but to estimate the coefficients with methods that account for the discrete nature of the data (Allison, 1982). However, as the model of Singer and Willett comes down to estimating a logistic regression model, it is more comprehensible and easy to apply.

Scott and Kennedy (2005) bring another improvement to the model of Singer and Willett with a view to using it for education research. In the former model there is only one event ("outcome") of interest, and when an individual leaves the dataset without experiencing the event in question he is censored. However, all ways of censoring must be independent of the outcome of interest. If this is not the case, the event that leads to exclusion from the dataset must be treated as an outcome of interest too. For example, suppose that the outcome of interest is degree completion. If an individual leaves the dataset because data collection ends before he completes his doctoral degree, he can be treated as censored, as there is no reason that this event affects his probability of Ph.D. completion later on. However, if an individual leaves the dataset because he drops out of doctoral study, he cannot be treated as censored. Indeed, as he dropped out, he cannot complete his degree anymore, which means that dropout is not independent of Ph.D. completion. The two events are called competing risks and both must be treated as outcomes of interest. Scott and Kennedy develop a discrete-time model that incorporates competing risks. In their model, the hazards of all events of interest are studied simultaneously, conditionally on the fact that the individual has not already experienced another outcome of interest.

The precise method used in this article is thus the discrete-time competing risks survival analysis presented in Scott and Kennedy (2005). This method is used in none of the articles reviewed in which time to Ph.D. completion or dropout is analyzed. In his study of time to the doctorate in Southern U.S. universities, Wilson (1965) uses descriptive statistics to compare characteristics of faster and slower groups in different disciplines. To predict time to degree attainment at UCLA, Abedi and Benkin (1987) use a stepwise multiple regression analysis. Baird (1990) mainly computes correlations between median time to the doctorate of a program and program characteristics in 228 U.S. universities. Stricker (1994) also uses correlations between departmental median times and institutional characteristics in American universities, but partials out a set of student characteristics. Finally, Seagram *et al.* (1998) analyze data on speed of degree completion at York University using chi-square analysis, analysis of variance and a multiple linear regression. All these authors avoid the censoring problem by only analyzing data on those students who had already completed their degree. However, as Willett and Singer (1991) note, they underestimate the times to degree by doing so. Some authors use survival methods, but not the discrete-time competing risks one of Scott and Kennedy. In their article on Ph.D. attainment at Flemish universities, Visser *et al.* (2007), use Cox's continuous-time hazard model with one outcome of interest. Siegfried and Stock (2001) also incorrectly use a continuous-time survival method with one outcome of interest in their study of time to degree in U.S. economics departments. Booth and Satchell (1995) use a continuous-time hazard model in their U.K. study, but incorporate competing risks. Ehrenberg and Mavros (1995) and van Ours and Ridder (2003) take the discrete aspect of their data into account in their competing risks duration models respectively at Cornell University and at three Dutch universities. However, they do this by discretising the continuous-time model. As said previously, this is less intuitive and comprehensible than the method of Scott and Kennedy, which can be estimated with a multinomial logistic regression program. While we did not find any analysis of the doctoral process in which the method of Scott and Kennedy is used, Arias Ortiz and Dehon (2011) apply this method in their study of degree completion and dropout in lower university levels at the ULB.

Although not all authors mentioned in this section focus on the "when" question or apply a (right) survival method when they do, their results are still interesting to comment, as the associations found between a variable and an outcome can guide our own reflection about the determinants of times to dropout and to degree. In the following, we will outline the results found in previous research concerning the five explanatory variables that are most interesting for our research: field of study, type of financing received, ability, gender and nationality.

The variable that is probably most often related to success or failure at the doctoral level is the field of study in which the student pursues his Ph.D. It is noteworthy that all articles that take the area of research into account find significant results for this variable. In general, disciplines in natural sciences, engineering and life sciences are associated with higher completion rates, lower dropout rates and lower times to degree than disciplines in social sciences and humanities. For example, Seagram *et al.* (1998) conclude that graduates in the natural sciences complete their studies significantly faster than students in the humanities and social sciences. Wilson (1965) finds the same result in his study of time to the doctorate in Southern U.S. universities. Visser *et al.* (2007) find large differences in the times to degree across disciplines, with a relatively short duration in sciences and a relatively long duration in law and social and economic sciences. Abedi and Benkin (1987), Baird (1990), Booth and Satchell (1995), Mooney (1968), Ott *et al.* (1984) and Park (2005) also find differences between fields of study.

A second variable that usually has an influence on the doctoral process is the type of financial support that the Ph.D. candidate receives. Students with grants and research assistantships often have a more successful doctoral path than teaching assistants and students who finance doctoral study with a job outside university or own earnings. For example, Abedi and Benkin (1987) find that source of support is the most important variable in predicting time to degree. In particular, a doctoral student who supports himself and his family through off-campus employment takes longer than the mean time to complete his doctorate. Ehrenberg and Mavros (1995) conclude that, in general, fellowships and research assistantships increase completion probabilities and decrease dropout probabilities relative to teaching assistantships and to all other forms of financing, such as loans and self-support. Remarkably, the effect of financial support is mostly felt through its influence on completion and dropout *rates*, rather than on the *time* to dropout or Ph.D. attainment. Stock *et al.* (2006) find that research assistantships are associated with lower two-year attrition than teaching assistantships in economics, once other factors are controlled for (including student ability). Seagram *et al.* (1998), Siegfried and Stock (2001), Stricker (1994), Visser *et al.* (2007) and Wilson (1965) also find differences associated to the type of financial support. Interestingly, Booth and Satchell (1995) conclude that British Ph.D. candidates with financial support from a research council or other funding body do not have significantly different degree and dropout rates than other students once other variables are controlled for.

A lot of authors include some measure of ability of the student in their analyses. In these cases, the ability variable is often related to an increased completion probability and a lower dropout probability, and with a lower time to degree. For example, both of the “academic achievement” variables that Mooney (1968) uses (whether the student was elected to Phi Beta Kappa and the selectivity of the college attended), increase significantly the probability of success for males and females. Visser *et al.* (2007) find that the result of the Master degree affects the Ph.D. attainment rate. The ability measures used by Grove *et al.* (2007, verbal and quantitative Graduate Record Examination scores), Stock *et al.* (2006, GRE scores) and van Ours and Ridder (2003, time taken to finish the undergraduate degree) also have an influence on success or failure at the doctoral level. However, Ehrenberg and Mavros (1995) find that their measures of student ability (verbal and mathematics GRE scores) only have an impact in physics and economics. Lunneborg and Lunneborg (1973) conclude that undergraduate GPA, letters of recommendation and quality of undergraduate department are correlated positively with first-year performance, but not with eventual success.

In their research on doctoral education, authors have found mixed results concerning gender. On one hand, a majority of authors find that gender is related to completion rate, attrition rate, or time to degree attainment or dropout (Abedi and Benkin (1987), Booth and Satchell (1995), Grove *et al.* (2007), Lunneborg and Lunneborg (1973), Mooney (1968), Stricker (1994), van Ours and Ridder (2003) and Visser *et al.* (2007)). However, not all of these studies control for other variables. In the studies that do, there are some interesting results. Abedi and Benkin (1987) conclude that gender is one of the most important variables to predict time to degree at UCLA, with men taking on average 1.2 years less than women to complete their doctorate. Van Ours and Ridder (2003) find that, in The Netherlands, females have higher attrition rates. Finally, Visser *et al.* (2007) conclude that success rates are higher for males than for females, even when other factors are taken into account. On the other hand, Park (2005), Seagram *et al.* (1998), Siegfried and Stock (2001) and Stock *et al.* (2006) conclude that gender has no influence on attrition or time to completion of the doctoral degree. Ehrenberg and Mavros (1995) and Ott *et al.* (1984) find that gender only has an impact in some disciplines.

Finally, the influence of nationality or citizenship found in previous research is also noteworthy. When this variable has significant results, the citizens of the country always have *lower* completion rates, *higher* dropout rates and/or *higher* times to degree than foreigners (Ehrenberg and Mavros, 2005; Park, 2005; Siegfried and Stock, 2001; Stock *et al.*, 2006 (only for women)). Grove *et al.* (2007) and Ott *et al.* (1984) find no significant influence of citizenship.

Other student-level variables that have an influence on the doctoral process are age (Grove *et al.*, 2007; Lunneborg and Lunneborg, 1973; Park, 2005), number of dependents or children (Abedi and Benkin, 1987; Lunneborg and Lunneborg, 1973; Siegfried and Stock, 2001 (only for women); Wilson, 1965), marital status (Lunneborg and Lunneborg, 1973), undergraduate university (Park, 2005; van Ours and Ridder, 2003; Wilson, 1965 (only for biosciences and physical sciences)), field of study of undergraduate degree (Lunneborg and Lunneborg, 1973; Wilson, 1965), supervisor and supervisory experience (Seagram *et al.*, 1998; van Ours and Ridder, 2003), taking a job prior to completing the Ph.D. (Siegfried and Stock, 2001; Wilson, 1965) and postdoctoral plans (Abedi and Benkin, 1987; Baird, 1990).

Finally, Baird (1990) and Stricker (1994) put a special emphasis on departmental and university-level characteristics, arguing that they are at least as important as student-level variables. Other authors also find departmental characteristics to be significant determinants of doctoral success or failure. Such variables include, among others, quality of the program (Baird, 1990; Stock *et al.*, 2006), size of the department or the graduate school (Mooney, 1968; Stricker, 1994 (only for psychology)), student-faculty ratio (Stricker, 1994 (only for psychology)) and departmental climate and integration of the students (Stock *et al.*, 2006 (only for women)).

3. The doctoral system of the French Community of Belgium

A student who wants to start a doctorate in one of the universities of the French Community of Belgium must have successfully completed a second cycle of higher education studies (as defined by the Bologna Process, often called “Master” degree) or a similar degree, and must be considered able to carry out scientific research. Moreover, he must find a thesis supervisor and a research unit of a university department that is willing to take him on. Finally, the department in which he plans to pursue his doctoral education must generally accept the subject of his Ph.D. thesis. Admission commissions of universities often decide that students who completed their second cycle studies at least “*cum laude*” (*i.e.* with an average of at least 70%) are automatically admitted to doctoral study provided they found a research unit and a thesis supervisor. Students who completed their second cycle “without honors” (*i.e.* with an average of less than 70%) must then have a special authorization to start doctoral study. This is the case at the ULB.

In the FCB, the time dedicated to working on a Ph.D. thesis corresponds to at least 180 ECTS credits[8]. Furthermore, the system includes a doctoral training of 60 ECTS credits. As one academic year corresponds to 60 ECTS credits, the time taken to attain the doctoral degree is in principle of at least four years, although a lot of students need more time to complete their Ph.D.

There exist different types of financing for Ph.D. candidates, which differ in terms of source, nature of duties, maximum length of time during which they are awarded, selectivity and amount awarded. The most common types of support for doctoral education in the FCB are presented in Table 1.

INSERT Table 1: Presentation of the most common types of financing for the pursuit of a doctorate in the French Community of Belgium

The FNRS allocates two types of grants for *maximum* four years: FNRS doctoral fellowships and FRIA doctoral fellowships. The FRIA fellowship is only accessible to doctoral students in science and technology or health sciences. Although both types of support are highly selective, the FNRS doctoral fellowship is more oriented on results obtained in second cycle studies than the FRIA doctoral fellowship. A lot of students also pursue a doctorate with an “ULB grant”, which generally also lasts for *maximum* four years. The funds for these grants come mainly from donations to the ULB or research contracts between the ULB and a firm or government. Finally, an assistantship position is also a common type of financing for doctoral education in the FCB. Contrary to students benefiting from the three grants presented earlier, an assistant dedicates

[8] In the Bologna Process, an ECTS (“European Credit Transfer and Accumulation System”) credit corresponds to 25 to 30 hours of work.

50% of his time to research and the other 50% essentially to teaching duties. Assistantships last for *maximum* six years in full-time equivalent. ULB grants and assistantships are less selective than the two grants from the FNRS. Note that some students also pursue a doctorate with a researcher contract as financing[9]. It is important to keep in mind that the financing of a student can vary in the course of doctoral study. Statistics about doctoral financing at the ULB can be found in Table 2. Besides students who really receive no support, the category "no financing" includes all types of financing that the administration of the ULB cannot detect, such as financings from governments of other countries or from external organizations that do not arrive at the ULB. It is interesting to note that 44% of all students have "no financing", 60% of which are foreigners.

INSERT Table 2: Descriptive statistics for the whole sample based on number of individuals (n=3,092)

4. Data

The sample used in this study consists of students who pursued a doctorate at the ULB, the *Université libre de Bruxelles*. The data cover the period from the academic year 2001-2002 to the academic year 2009-2010[10]. The dependent variable under study is the outcome of interest experienced in a particular time period, *i.e.* dropout of doctoral study or degree completion. If none of these outcomes are experienced in the period, the nonoutcome 0 occurs. Ph.D. completion occurs when the student defended his thesis during the time period. Although the precise date of thesis defense was available, the time period used in this study is the academic year. The main reason for this is that dropout could only be observed once a year[11]. Indeed, dropout was detected when an individual left the database – and thus did not register anymore – without defending his thesis. A student is considered to have experienced the dropout outcome during the last year in which he was registered. For example, if a student registers for doctoral study during five years and does not enroll for a sixth year without having defended his thesis, he experiences the outcome dropout during his fifth year of study.

The analysis time is the number of academic years during which a student enrolled for doctoral study (without gaps in the registration period), as the database from the ULB only contains information about doctoral students that actually registered in a particular year. Since the

[9] Two main types of students are in this case. On one hand, researchers at the ULB can decide to pursue a doctorate on top of the research they carry out as part of their contract. On the other hand, researcher contracts are sometimes given to students who cannot benefit from doctoral grants anymore or to students with a part-time assistantship. In this case, they are financed specifically to pursue a doctorate. As the two types of students could not be distinguished and only 5% of doctoral candidates have researcher contracts, we did not focus on this type of financing in this article.
[10] At the ULB, an academic year starts on the 15th of September and ends on the 14th of September.
[11] A second reason is that the (time-varying) explanatory variables also vary per academic year in the database.

academic year 2000-2001, doctoral students are compelled to enroll each year that they spend working on their thesis. Previously, a lot of students did not register every year, although they worked on their thesis, and some only registered when they wanted to defend their thesis. As some time has been necessary to adapt to the new legislation, there are still individuals who do not enroll every year. Although this phenomenon is probably fading away with years, it still implies that we do not have data for the whole period of doctoral study for some students. The final dataset used for the analyses consists of 3,092 individuals[12]. It contains one line per year that an individual was registered, or 10,235 "person-year" lines.

The explanatory variables that will be used to estimate the hazards of the outcomes, apart from dummies indicating the year under study, can be classified into two categories. The first set of variables is available for all individuals in the database. Analyses with these predictors will in consequence be based on the whole sample (n = 3,092[13]). These variables are the following:

- Gender (time-invariant): dummy variable **Female** that equals 1 if the individual is female.
- Nationality (time-invariant): dummy variable **Belgian** that equals 1 if the individual is Belgian.
- University of second cycle (time-invariant): dummy variable **ULB** that equals 1 if the university in which the second cycle of studies was completed is the ULB.
- Type of financing (time-varying): categorical variable **Financing** with seven categories: assistantship, FNRS doctoral fellowship, FRIA doctoral fellowship, ULB grant, researcher contract, other financing and no financing. The financing variable is constructed as follows: for each year, the category taken on by the variable is the type of financing that the student received most until then[14]. For example, if a student was an FNRS doctoral fellow during four years and then received no financing – or another type of financing – for two years, the financing variable still takes on the value "FNRS doctoral fellowship" during these two years. The variable was constructed as such to take into account the fact that the history of financing matters. Indeed, if the financing variable was merely equal to the type of financing received during each year, the student mentioned above would have "no financing" during years five and six. However, it is probable that the FNRS grant from which the student benefited still has an impact on the hazards of dropout and Ph.D. completion in the last two years.

[12] 158 students with gaps in the registration period were excluded from our analysis.
[13] n refers to the number of individuals, whereas N refers to the number of person-year lines.
[14] If a student received two or more types of financing during an equal number of years, one of these financings was selected based on the following order: FNRS doctoral fellowship, FRIA doctoral fellowship, ULB grant, assistantship, researcher contract, other type of financing and no financing.

- Field of study (time-invariant): categorical variable **FOS** with three categories: human sciences, health sciences, and science and technology. Each broad field of study is a grouping of several disciplines (see Table 12 in Appendix).
- Age at start of Ph.D. (time-invariant): dummy variable **Age** that equals 1 if the individual started his Ph.D. at 26 or older[15].
- Marital status (time-varying): categorical variable **MS** with five categories: single, married, divorced, separated and widowed. As only 1.29% of the person-year lines in the database had divorced, separated or widowed as marital status, these categories were added to the single class to form the category "not married". The resulting variable is **MSd**, a dummy variable that equals 0 if an individual is not married in a particular year and 1 if he is.

The second set of explanatory variables concern the second cycle of study completed before the start of the doctoral process. These covariates were not available or could not be computed for all students in the database. Most individuals for whom data were missing completed their second cycle in a foreign university. Indeed, 73% of the students with no missing data pursued their second cycle studies in Belgium, whereas this percentage was of only 19% in the group of individuals for whom data were missing. Therefore, all analyses containing these "second cycle explanatory variables" will be based on a subsample, called the "Belgian university subsample", consisting of students who pursued their second cycle in a Belgian university and for whom no variables were missing (n = 1,915). The second cycle variables are the following:

- Result obtained in second cycle studies (time-invariant): categorical variable **Result** with three categories: *summa cum laude*, *magna cum laude*, and *cum laude* or success without honors. In Belgium, a university degree is awarded *cum laude* if a 70% average is achieved, *magna cum laude* for an 80% average and *summa cum laude* for a 90% average. When an average between 60% and 70% is achieved, the student receives his university degree, but without honors. In this case, he has no access to doctoral study, unless he receives a special authorization. Therefore, only 80 out of the 1,915 individuals in the Belgian university subsample obtained their second cycle degree without honors. These students were gathered with the ones who completed their second cycle *cum laude*.
- Equivalence between discipline of second cycle and discipline of doctorate (time-invariant): dummy variable **Equivalence** that was constructed based on the disciplines of the second cycle and the doctorate in the database, and that equals 1 if the second

[15] 26 is the median age at start of Ph.D. in the sample.

cycle and the doctorate are pursued in the same discipline or in disciplines between which it is easy to switch.

- Failure in second cycle (time-invariant): dummy variable **Fail** that was constructed based on the number of years taken to complete the second cycle studies. It equals 1 if the student failed at least one year during his second cycle.

Table 2 contains summary statistics about the whole sample (n = 3,092). The sample consists for 55% of men, and the ratio of Belgian students to foreign students is of three to two. The financing statistics are based on the type of financing that the individual received most during his whole period of registration for doctoral study. We can see that the most frequent type of financing was a grant from the ULB (14%), followed by a FRIA doctoral fellowship (13%), an assistantship (11%), an FNRS doctoral fellowship (9%) and a researcher contract (5%). 44% of the students in the sample were *mostly* unfinanced during doctoral study. Note that the percentage of students who did not receive any type of financing during the *whole* period of doctoral education is of 41%. 41% of the students in the sample pursued a doctorate in human sciences, 38% in science and technology and 21% in health sciences. We can already see interesting patterns in the proportions of individuals who experienced dropout or Ph.D. completion in the subsamples. For example, there are important differences between the fields of study, with science and technology having the highest completion rates and the lowest dropout rates, followed by health sciences. The types of financing also have very different dropout and Ph.D. proportions. For example, almost 50% of those who were mostly unfinanced during doctoral study dropped out, while this proportion was of only 10% for the FNRS doctoral fellows and 12% for the FRIA students. For the teaching assistants, the dropout rate was of 27%. 42% of those who were mostly financed by a FRIA doctoral fellowship and 38% of the FNRS doctoral fellows attained their doctorate, but only 25% of teaching assistants and 19% of the unfinanced students did.

Table 3 shows descriptive statistics for the Belgian university subsample (n = 1,915). Most students did not fail a year of their second cycle studies and chose a discipline for their Ph.D. that was comparable to the one of their previous studies. 16% of the students obtained their second cycle degree *summa cum laude*, 50% *magna cum laude* and 30% *cum laude*[16]. For the second cycle variables we can also see interesting patterns. For example, as the result obtained in the second cycle of study becomes better, the proportion of students who finished their Ph.D. increases and the proportion of students who dropped out decreases. There are also big differences in dropout and Ph.D. proportions between the subsample of students who did not

[16] In comparison, of all students who finished their second cycle at the ULB during the period 2000-2001 until 2009-2010 (also those who did not start a doctorate), 4% obtained their degree *summa cum laude*, 25% *magna cum laude*, 48% *cum laude* and 24% without honors.

fail a year during their second cycle studies and the subsample of the ones who did. Table 13 in Appendix contains the summary statistics of all the other variables for the Belgian university subsample. We can for example see that the proportion of unfinanced students is of only 29%, while the proportion of FNRS doctoral fellows, students with FRIA doctoral fellowship and teaching assistants is higher than in the whole sample. The proportion of Belgian students is of 88%, compared to 58% in the whole sample.

INSERT Table 3: Descriptive statistics for the Belgian university subsample based on number of individuals, only "second cycle variables"(n=1,915)

5. Exploratory analysis

In this descriptive analysis, the hazards of the outcomes dropout and degree completion for each year after the start of registration for doctoral study are computed. As the data cover the period 2001-2002 until 2009-2010, the last year for which the hazard can be modeled is T=9 (ninth year after the start of doctoral study). At this point, it is useful to outline two essential concepts in survival analysis[17]. The **risk set in a particular year t (t = 1, 9)** ("those at risk") is the group of students in the sample who we know can experience one of both outcomes during t. The risk set does not include individuals who have already experienced one of both events or who were censored in one of the previous years. The **hazard of outcome k (k=1,2) in year t (t=1, 9)** [$\boldsymbol{h(k,t)}$] is the probability that a randomly selected individual experiences k (dropout or degree completion) in year t, given that he has experienced the nonoutcome 0 in every year before t. Thus, discrete-time hazard is a conditional probability. The sample hazard probability of outcome k in year t can be obtained by dividing the number of students who experienced outcome k in year t by the number of students in the risk set in the beginning of year t. As the sample hazard is based on the risk set, it can be computed in every year. The maximum likelihood estimates for the hazard probabilities $\hat{h}(k,t)$ (k=1,2 and t=1, 9) correspond to the sample hazard probabilities. These probabilities can be found for the whole sample (n = 3,092) in Table 4.

INSERT Table 4: Presentation of the survival statistics for the whole sample (n=3,092)

We can see that of the 3,092 individuals, 389 dropped out and 21 completed their Ph.D. during year one. This means that, in year one, the estimated hazard of dropout is 12.6%, while this hazard is of 0.7% for Ph.D. completion. Moreover, 313 individuals were censored at the end of year one. As a consequence, the risk set for year two is composed of 2,369 students. As 263 individuals dropped out and 54 completed their degree in year two, the sample hazard

[17] The following definitions are based on the ones of Scott and Kennedy (2005) and of Singer and Willett (1993).

probabilities are respectively of 11.1% and 2.3% in this year. This reasoning can be applied to all subsequent years.

To have an idea of the evolution of the hazard of an outcome over time, we can plot the sample hazard function for the outcome, which is the sequence of hazard probabilities across the nine years (Figure 1).

INSERT Figure 1: Maximum likelihood estimation of the hazard functions (whole sample)

We can see that in the first years, the hazard probabilities for the outcome dropout are higher than the ones for the outcome Ph.D. completion. This is logical, as completing a doctorate normally takes at least four years. However, 149 of the 3,092 individuals in the sample finish their Ph.D. in one, two or three years after the start of registration for doctoral study. This phenomenon is probably due to the fact, discussed above, that they were already working on their Ph.D. before registering for the first time. A substantial group of students (175) complete their doctorate in four years, but the hazard of Ph.D. completion really jumps in year five, to more than 30%. Table 4 shows that the majority of students who complete their doctorate do so in year five (275). After year five, the hazard of Ph.D. attainment increases slightly each year, except in year eight. However, after year seven, the hazard estimates must be interpreted with caution. Indeed, they are based on a small number of observations, as only 69 students survived after year seven and 15 after year eight. The estimated hazard of dropout remains fairly stable along the doctoral process around 10% (the estimations for years eight and nine are again based on very few observations).

While the hazard functions are useful to have an idea of the probability of experiencing an outcome in one single year, it is also interesting to analyze the probability of having experienced an event *by* a certain year. Therefore, cumulative hazard probabilities and their functions over time are needed. The **cumulative probability** $M(t)$ **(*t*=1,..., 9)** is the probability that any of both outcomes occurs in the first t years. The sample cumulative probability can be computed with the following formulas:

$$M(1) = h(k = 1,1) + h(k = 2,1)$$

$$M(t) = [h(k = 1, t) + h(k = 2, t)] \times [1 - M(t - 1)] + M(t - 1) \; for \; t > 1. \quad (1)$$

The sequence of estimates $\widehat{M}(t)$ across the nine years gives the estimated cumulative probability profile for both outcomes (Figure 2, right-hand graph). It is interesting to look at the median survival time of an individual in the sample, which is when 50% of the students in the sample experienced one of both events. In this case, the median survival time is slightly more than four years. We can also see that more than 90% of the sample experienced one of both outcomes by year seven. Knowing that in addition the sample contains very few observations for the years eight and nine, we decided to gather for all following analyses the years seven, eight and nine.

INSERT Figure 2: Maximum likelihood estimation of the cumulative hazard functions (whole sample)

The **cumulative probability of *k* (*k*=1,2) in year *t* (*t*=1,...,9)** [$\boldsymbol{M(k,t)}$] is the probability that outcome k occurs in the first t years. The sample cumulative hazard probabilities (Figure 2, left-hand graph) can be computed with the following formulas (k=1,2 and t=1,...9):

$$M(k,1) = h(k,1)$$

$$M(k,t) = h(k,t) \times [1 - M(t-1)] + M(k,t-1) \; for \; t > 1. \qquad (2)$$

Note that before year six, the cumulative probability of dropout is higher than the cumulative probability of Ph.D. completion.

In order to get a first idea of the influence of explanatory variables on the hazard profile of the two outcomes, we can compute the hazard and cumulative hazard functions for subsamples of individuals with different values of explanatory variables. However, it is important to note that the impact of the explanatory variables that can be seen on these graphs can be due to differences in other predictors, as this is only a bivariate analysis. In Figure 3 we computed the hazard functions of dropping out and completing the degree separately for the subsample of males and the subsample of females. Looking at the sample hazards, it seems that gender has no influence on dropout or on Ph.D. completion[18].

INSERT Figure 3: Hazard functions by gender (whole sample)

Figure 4 shows the cumulative hazard functions for the three subsamples based on field of study. The field of study in which the doctorate is pursued seems to have an influence on both dropout and Ph.D. completion. Students in human sciences have the highest dropout probabilities, followed by health sciences and science and technology. The result is that by year seven more than 50% of the human science students have dropped out, while this percentage is of about 40% for health sciences and about 30% for science and technology. Concerning

[18] The sharp drop in Ph.D. completion sample hazard in year six for women must probably be attributed to the relatively small number of observations, as only 192 women survived until year six in the sample. 55 of them obtained their doctorate in year six.

completion hazards, the influence of field of study is important from year five onwards. Indeed, for science and technology the cumulative completion probability jumps to more than 40% in year five, while it is of about 25% for health sciences and human sciences. In years six and seven the cumulative probabilities of health sciences exceed the ones of human sciences, but the difference is much smaller than the one with science and technology. The fact that the dropout probabilities are higher for human sciences and the completion probabilities lower concord with the general idea that students who pursue a doctorate in this field of study complete less often their doctoral education and take a longer time if they do so. Moreover, it is in line with the results of previous research on doctoral education. However, this finding will be moderated later on, when we control for other variables.

INSERT Figure 4: Cumulative hazard functions by field of study (whole sample)

For the variable type of financing, the cumulative hazard probabilities also display the clearest patterns (Figure 5). Remember that the financing variable is each year equal to the category of financing that was most received until then, taking thereby the history of financing into account. The graph of dropout hazards reveals three clear groups of cumulative probabilities. Students who have no financing during most of their doctoral education are the ones who drop out most. This group is followed by the individuals with ULB grants and assistantships. Finally, FNRS doctoral fellows and students with FRIA doctoral fellowship drop out least. The type of financing received also influences Ph.D. attainment hazards, where we can see important differences from year four onwards. FNRS doctoral fellows have the highest cumulative completion hazards, followed by FRIA students, individuals with ULB grants, teaching assistants and unfinanced students. It is interesting to note that the jump in cumulative completion hazard in year four is only present for FNRS doctoral fellows. In subsequent years, the difference with FRIA students becomes smaller.

INSERT Figure 5: Cumulative hazard functions by type of financing (whole sample)

The variable **Age** (Figure 6) also influences the hazards, but here the interpretation of the graph must be made with caution. Two important factors explain the pattern that can be seen on the graph of the hazard profiles. On one hand, some individuals do not enroll at university for the first year(s) that they work on their Ph.D., which implies that they are older when they register for the first time and that the hazards of getting a Ph.D. in the first years are higher for them (looking at years one, two and three, which are years in which it is very unusual to already have finished a doctorate). Including the variable **Age** in the model presented later will thus be a way of controlling for the problem of non-registration during the first years of doctoral study. On the other hand, it seems that people who start their Ph.D. at 26 or older tend to have lower Ph.D. completion hazards (when looking at the years four or more, which are the years in which a Ph.D. can in general be completed if the individual registered since the beginning of doctoral study) and higher dropout hazards.

INSERT Figure 6: Hazard functions by age at start of Ph.D. (whole sample, 26 is the median age at start of Ph.D. in the sample)

The two following graphs concern second cycle variables and are based on the subsample of students who pursued their second cycle studies in Belgium and for whom no data were missing ("Belgian university subsample"). Both cumulative dropout and Ph.D. completion hazards show important differences according to the result obtained in the second cycle studies (Figure 7). Cumulative dropout hazards are always lowest for students who completed their second cycle *summa cum laude*, followed by *magna cum laude*, and *cum laude* or without honors. The cumulative Ph.D. completion hazards start to be different from year four onwards, as before it is unusual to complete a doctorate. Again, *summa cum laude* students have the highest Ph.D. hazards, followed by *magna cum laude* students, and *cum laude* students or students who completed their second cycle without honors. In conclusion, the **Result** variable seems to have an important influence on completion and dropout hazards. However, we have to remember that this is only a bivariate analysis, and that the real impact of this variable can only be detected when controlling for other variables, such as financing. In particular, there is a link between type of financing received and the result obtained in second cycle studies. The question thus arises whether it is the type of financing received or the result obtained in second cycle that drives the influences of those two variables on the dropout and completion hazards.

INSERT Figure 7: Cumulative hazard functions by result obtained in second cycle studies (Belgian university subsample)

Finally, the cumulative hazard profiles of the variable **Fail** (Figure 8) indicate that students who failed at least one year during their second cycle studies have consistently higher cumulative dropout rates than those who finished their second cycle studies on time. Moreover, they also have lower cumulative completion probabilities from year four onwards.

INSERT Figure 8: Cumulative hazard functions by failure in second cycle (Belgian university subsample)

6. Model-based approach

6.1 Introduction

In order to control for multiple variables when analyzing the impact of predictors on the hazard of one of both outcomes, a model-based approach is necessary. Cox (1972) suggested that, given that the elements $h(k,t)$ are probabilities, they can be modeled using a logistic dependence over the predictors. In consequence, Singer and Willett (1993) prove that in a discrete-time setting with only one outcome of interest a traditional logistic regression can be used to obtain maximum likelihood estimates of the coefficients. Scott and Kennedy (2005) show that for discrete-time competing risks models, a multinomial logistic regression program can be used. Applied to this case, the specification of the hazard in a competing risks setting states that the hazard for subject *i* (*i*=1,…,n) of outcome *k* (*k*=1,2 – dropout or Ph.D. completion) at time *t* (*t*=1,…,9) is modeled as:

$$h_i(k,t) = \frac{\exp\,(X_{it}\beta_k)}{1+[exp(X_{it}\beta_{k=1})+exp(X_{it}\beta_{k=2})]} \tag{3}$$

where X_{it} is a row vector of covariate values for subject *i* at time *t* of length *P* and β_k is a column vector of *P* parameter estimates for outcome *k*.

Taking logistic transformations of both sides of the previous equation leads to:

$$\ln\left(\frac{h_i(k,t)}{h_i(0,t)}\right) = X_{it}\beta_k \tag{4}$$

where $h_i(0,t)$ is the hazard of the nonevent for individual *i*.

We can see that we assume a linear relation between the explanatory variables and the “logit-hazard”, the natural logarithm of the outcome-specific hazard ratio. The outcome-specific hazard ratio is the ratio of the hazard of outcome *k* (*k*=1,2) to the hazard of the nonoutcome 0. When applying the exponential function to both sides of Equation 4, we see that this hazard ratio is multiplicative in the exponentiated covariates. The coefficients can be estimated with a standard multinomial logistic regression method where all the person-year lines observed in the sample are treated as independent observations. The dependent variable is then a categorical variable that accounts for the outcome experienced by the individual during the year (no outcome, dropout or Ph.D. completion) and the explanatory variables include dummies for the years.

6.2 Baseline profile of risk over time

To analyze the effect of time as only predictor of the hazard, we put solely time dummies as covariates in the model. This is referred to as the baseline profile of risk over time. The time dummy D_{it} take the value 1 when the observation for subject i comes from year t and 0 otherwise. They are defined for all periods until T, the last time period observed for anyone in the sample. Although the last time period in the dataset is year nine, the years seven, eight and nine will be grouped in the model to have more observations for this period. By using seven mutually exclusive time dummies, it is not necessary to include a constant in the model. In fact, the coefficients of the time dummies act as multiple constants (one per year). The sequence of parameter estimates $\hat{\beta}_{kt}$ tied to the different time dummies represent the shape over time of the logit-hazard function for the whole population. When exponentiating them, we find the outcome-specific hazard ratios for each year. These can be converted into outcome-specific hazards, which are identical to the maximum likelihood estimates of the hazards presented in the explanatory analysis (Table 4). The coefficients, outcome-specific hazard ratios and outcome-specific hazards from the baseline model using the whole sample are reported in Table 5.

INSERT Table 5: Results of fitting the baseline model using the whole sample

6.3 Presentation of the full models and assumptions

To the baseline profile of risk, we can add other predictors. The coefficient $\hat{\beta}_{kq}$ associated to the covariate X_q estimates the shift of the sequence of outcome-specific logit-hazards due to a change in the covariate, while controlling for other variables. Time-varying covariates, *i.e.* financing and marital status in this case, can lead to different shifts of the logit-hazard profile of an individual in different years.

Before adding explanatory variables to the model, the three assumptions underlying the discrete-time hazard model must be verified[19]. First, for the linear-logistic model to be a valid representation of reality, equal differences in a predictor must have an equal impact on the logit-hazard. This assumption is called the **linearity assumption** and is equivalent to the one made in linear regression analysis. It can be checked through exploratory data analysis or statistical inference (Singer and Willett, 1993) and is only relevant for non-categorical variables.

The second assumption concerns **no unobserved heterogeneity**. As there is no error term in the hazard model, all variation across individuals must be accounted for by the explanatory variables included in the model. No satisfying method has yet been found to deal with the

[19] The three assumptions are presented in more detail in Singer and Willett (1993).

problem of unobserved heterogeneity (Allison, 1982). The assumption will therefore be ignored here.

Finally, the **proportional hazards assumption**, which underlies all survival methods, must be checked. The model presented until here is a main effects model, as a covariate has the same effect on the logit-hazard whatever the time period. This means that the whole logit-hazard function will be shifted vertically when the covariate takes another value. For example, a male and a female with the same values for the other predictors will have the same shape of logit-hazard, but at another level. The difference between the logit-hazards of both students is the same in each year. Time homogeneity can be checked easily with graphical methods: logit-hazard profiles estimated separately for subsamples based on different values of a covariate must be approximately parallel. If this is not the case, interactions between the predictor and the time dummies must be included in the model, allowing the effect of the predictor to vary over time. The model is not a main effects model anymore, but an interaction model (Singer and Willett, 1993). Wald tests of equality of the coefficients attached to the different interaction terms of a covariate can then confirm the interaction detected graphically. Indeed, the graphical analysis provides only a partial picture, as it does not control for other variables. If the null hypothesis of equality of coefficients is rejected, the impact of the covariate evolves over time and the interactions with time must be kept in the final model. After graphical analysis and Wald tests for the equality of coefficients, it appeared that the effect of field of study evolves through time for both categories (health sciences and science and technology), at least for one of both outcomes. Therefore, interaction terms between **FOS** and time must be included in the full model. The other variables for which interaction terms must also be included are **Financing** for the categories FNRS doctoral fellow, FRIA doctoral fellow and ULB grant[20], **Aged**, **MSd** and **Result**[21].

After the verification of model's hypothesis, we mainly estimated two models with additional predictors. Besides the time dummies, the first full model contains the predictors **Female**, **Belgian**, **ULB**, **Financing** (interacted with time for the categories FNRS doctoral fellow, FRIA doctoral fellow and ULB grant), **FOS** (interacted with time), **Aged** (interacted with time) and **MSd**. The control categories for **Financing** and **FOS** are assistantship and human sciences. The first model was estimated using the whole sample and the results of this estimation are given in Table 6. In addition to all the variables of the first model, the "second cycle covariates" **Result** (interacted with time), **Fail** and **Equivalence** were added in the second full model, which was estimated using the Belgian university subsample. The control category of the variable **Result** is

[20] Assistantship is the control category.
[21] *Summa cum laude* is the control category.

summa cum laude. Table 7 displays the results of fitting this second model. Except for the covariates **Financing** and **FOS**, the variables that were already included in the first model are not presented in Table 7. They can however be found in Appendix 5. The estimates presented in these tables must be interpreted as follows: the marginal effect of a covariate will be to increase the logit-hazard [ln $(\hat{h}_l(k,t)/\hat{h}_l(0,t)$] by an amount equal to the coefficient and to multiply the hazard ratio [ln $(\hat{h}_l(k,t)/\hat{h}_l(0,t)$] by the exponentiated coefficient. The marginal effect on the hazard $[\hat{h}_l(k,t)]$ depends however on the initial level of the hazard or, equivalently, on the values taken by all the other coefficients[22]. In order to have an idea of the magnitude of the effect of a variable on the hazard, it is thus necessary to select some sets of meaningful values for all variables[23]. A graphical presentation for one specific individual is useful in this respect. Concerning interaction terms, there is some discussion as to the interpretation of their effect in non-linear models (see for example Ai and Norton (2003), Buis (2010), Hoetker (2007) and Norton, Wang and Ai (2004)). In this model however, variables are only interacted with time dummies. Moreover, when an interaction term is included for a variable, the variable is not included on its own in the model anymore. This implies that these interaction terms are not like the ones we usually find in econometric models. Indeed, the idea here is to first specify a time period, with the corresponding time dummy acting as intercept. Once the time period has been chosen, there are no interacted variables anymore. In consequence, we can look, per year, at the coefficient and the marginal effect of all variables in the model to have an idea of their impact on dropout and Ph.D. hazards.

In the following, the most interesting results of the estimated models will be presented. We will first focus on the variables **Female**, **Belgian**, **ULB**, **Aged** and **Fail**. Next, we will discuss the impacts of type of financing and result of second cycle, which are linked. Finally, we will study the influence of field of study.

6.4 Gender, nationality, university of second cycle, age and failure in second cycle

The variables **Female**, **Belgian**, **ULB** and **Aged** were included in the first model, fitted using the whole sample (Table 6). First, it is interesting to note that neither gender nor nationality has a significant impact on either of both outcomes, once other factors are controlled for. Interestingly, having pursued second cycle studies at the ULB implies a significantly higher dropout hazard when we control for type of financing received. Although age at start of Ph.D. has

[22] Hoetker (2007) discusses the difficulties encountered when interpreting coefficients in logit and probit models and suggests some best practices.

[23] Alternatively, Train (1986) suggests to calculate the marginal effect for all observations and to average those effects. It is important to note that these average values can be very different from the marginal effect calculated for the "average" individual, *i.e.* when setting all explanatory variables at their mean.

almost no significant impact on dropout, the effect of **Aged** on Ph.D. completion is in line with what was discussed in the exploratory analysis. Indeed, in the third year the hazard is significantly higher for students who started their Ph.D. at 26 or older due to the non-registration effect. Note that for the first two years the estimates of the coefficients are not presented, as no student from the young group completed his Ph.D. in one year and only two in two years (compared to 21 and 52 respectively for the older students). In the second model estimated using the Belgian university subsample, the impact of these four variables is very similar.

The variable **Fail** was only included in the second model (Table 7). It has a significant impact, both on dropout and on Ph.D. completion. Indeed, students who failed at least one year during their second cycle studies have higher hazards of attrition and lower hazards of attaining the doctorate than their colleagues.

INSERT Table 6: Results of fitting the first model using the whole sample

INSERT Table 7: Results of fitting the second model using the Belgian university subsample (Selected variables)

The explanatory analyses and the results based on the model can be summarized in the following way. First, all analysis indicate that gender has no influence on dropout or Ph.D. completion, except for FNRS doctoral fellows in terms of dropout and teaching assistants in terms of Ph.D. completion. This result conflicts with the assumption of de Henau and Meulders (2003) that women withdraw more from doctoral study than men in the FCB. In addition, knowing that there are almost as many women as men enrolled at the doctoral level at the ULB, would also seem to suggest that there is no significant gender difference at this educational level in the FCB. This ties up with the idea of Meulders *et al.* (2005), who argue that differences in characteristics of men and women such as educational level do not explain the gender wage gap, while occupational segregation is much more important factors in the wage differences[24]. O'Dorchai (2009) shows that having children deepens the gender wage gap in most European countries. It would therefore have been interesting to analyze the influence of parenthood on doctoral dropout and completion. Moreover, de Henau and Meulders (2003) note that the age of motherhood often coincides with the period of thesis writing. Unfortunately, we did not have data on the number of children of doctoral students.

[24] It also emphasizes the question why the proportion of women in the scientific and academic staff of universities of the FCB is so low (see de Henau and Meulders, 2003).

Second, while nationality has no significant impact on the doctoral path, a student has a significantly higher dropout hazard if he pursued his second cycle at the ULB. By including the university variable, we wanted to account for the fact that some foreigners had already studied at the ULB previously. The fact that someone who pursued his previous studies at another university, which in almost 80% of cases is a foreign university, has lower dropout hazards is rather surprising, but is again reassuring. Moreover, it is in line with the finding of Ehrenberg and Mavros (1995) that U.S. citizen and permanent residents have larger dropout hazards than do foreign residents. This result might be due to the fact that pursuing a doctorate at the ULB is less straightforward for students who did not study at the ULB previously. Therefore, this might be a decision that they think through more seriously. If this is the case, it is important to encourage students who just finished their second cycle at the ULB to think their decision of doctoral study through and to make them aware of what the pursuit of a doctorate implies. Golde (2005) finds for example that attrition is often due to inaccurate expectations about the nature of doctoral education. Moreover, she argues that some students drop out because they perceive the job market to be poor. From the survey "Careers of Doctorate Holders" (CDH) in Belgium we can learn that 51% of the doctorate holders of less than 35 years are on temporary contract (Moortgat, 2011), as there are too few vacant permanent positions in universities. Moreover, although the qualities they gained thanks to their doctoral education are important, Ph.D. holders who turn to the private sector sometimes find that their degree is not valued by employers, who think they are overspecialized and individualistic (Durez et al., 2001). Making sure that the doctoral candidate is aware of these aspects of doctoral life is important to avoid dropout due to disenchantment. In this respect, Nelson and Lovitts (2001) argue that providing balanced information about the programs and preparing undergraduates for the culture of graduate school is essential to lower the attrition rate.

6.5 Type of financing and result obtained in second cycle of studies

When looking at the results of the first model (Table 6), the type of financing received seems to influence highly the doctoral path, both in terms of dropout and of Ph.D. completion. However, each type of financing is supposed to be attributed based on student skills, skills that we do not observe in our model. Thus, the effect that we capture could come from the fact that the model does not control for ability of the students. This means that it is not clear whether it is the fact of having a particular type of financing that enables some students to have low dropout and high completion hazards, or their "ability".

As the variable "results obtained in previous studies" is a major criterion for selection committees of all types of financing, it can be treated as a proxy for student skills or ability variable. However, **Result** is not a perfect proxy for the aspects that selection committees take into account for the allocation of financing. Increasingly, selection committees try to take into account other aspects, such as the research project. Indeed, having excellent results in previous studies certainly does not guarantee by itself that a student will be capable of completing a doctorate brilliantly as well. However, as "previous results" still play an important role in their decisions, particularly for grants like FNRS doctoral fellowships, this variable is probably closely related to ability as perceived by selection committees. Evidence suggests there is clearly a link between type of financing received and result of second cycle. To verify this, Table 8 displays the type of financing that the students received most during their doctoral education according to their result in the second cycle of study. Students who attained their second cycle degree *summa cum laude* have more easily access to FNRS doctoral fellowships and FRIA doctoral fellowships. Almost 70% of *summa cum laude* students have one of these two grants, while about 30% of *magna cum laude* students, 14% of *cum laude* students and 7% of students who finished their second cycle without honors do.

INSERT Table 8: Contingency table between Financing and Result based on number of individuals for the Belgian university subsample (n=1,915)

It is thus interesting to look at the estimation of the second full model, in which both **Financing** and **Result** are included (Table 7). We can see that the impacts of the different types of financing on dropout and Ph.D. completion hazards are similar to the ones of the first model. We will now look at dropout and degree attainment hazards separately.

Concerning dropout, the results are in line with the exploratory analysis. FNRS doctoral fellows drop out significantly less than teaching assistants in years one and three, and FRIA students in years one, three and four. With these grants, students can concentrate fully on their thesis, as they do not have teaching duties and they cannot take an outside job. Therefore, there is probably less chance that they feel discouraged after some years. Wald tests of equality of coefficients between FNRS doctoral fellows and FRIA doctoral fellows showed that the dropout coefficients of these two types of financing are not significantly different at usual confidence levels, except in year four, in which FRIA students have significantly lower dropout hazards than FNRS doctoral fellows. Students who are mostly unfinanced during their doctoral education drop out more than teaching assistants. This can be explained by the fact that assistants are much more integrated in the department thanks to their teaching duties (Nelson and Lovitts, 2001), integration being of high importance in doctoral education. Moreover, unfinanced students must often take a job outside university, which might be too much to handle after a while. Finally, unfinanced students also have nothing to lose when they drop out, which makes the decision of attrition probably easier. The differences in dropout might of course also be due to different levels of ability. For example, more than 50% of the students who are mostly unfinanced completed their second cycle degree without honors or *cum laude.* The type of financing received is thus dependent on the perceived ability of the student, which triggers the question of causality. We can see that the result obtained in the second cycle studies also has an impact on the dropout hazards. People who completed their second cycle degree without honors or *cum laude* drop out more in the first two years than students who completed their second cycle at least *summa cum laude*. Students who obtained a *magna cum laude* result drop out more than *summa cum laude* students, but only in year two. The fact that both **Financing** and **Result** are significant in terms of attrition hazards means that both variables have an impact on withdrawal, in addition to the influence of the information common to the variables. Thus, it is their ability *and* the fact of having a grant that enables FNRS doctoral fellows and FRIA students to have low dropout hazards. The influence of student ability on dropout is confirmed by the fact that there are significant differences between the coefficients of ULB grants, and the ones of FNRS doctoral fellowships or FRIA doctoral fellowships[25]. Indeed, although the three grants have similar characteristics, FNRS research fellowships and FRIA doctoral fellowships are more selective and thus given to the students who are considered most capable of pursuing research. Thus, financing, and particularly FNRS doctoral fellowships and FRIA doctoral fellowships, is allocated to students who will drop out less during their first years. In other words, the allocation of grants is efficient in the sense that it does not invest resources in students who will withdraw from doctoral study.

Turning now to Ph.D. completion, we can see that FNRS doctoral fellows and students with ULB grants have higher completion hazards in years four and five than teaching assistants. FRIA students also have higher Ph.D. attainment hazards, but only from year five onwards. All these types of financing are normally given during *maximum* four years. In comparison, assistantships are given for *maximum* six years, as they come with teaching duties. It is thus logical that these students finish their Ph.D.'s more often in four or five years than teaching assistants. However, it is important to stress that, while students with these types of grants have *higher* completion hazards than teaching assistants in years four and five, they do not have significantly *lower* completion hazards after year five. It is also interesting to note that, while FNRS doctoral fellows already have a higher completion hazard in year four, this is not the case for FRIA students. Wald tests of equality between the coefficients of these two types of financing confirmed that the fourth year is the only year in which the coefficients of FNRS doctoral fellows and FRIA students

[25] Wald tests of equality of coefficients indicate that FNRS doctoral fellows have significantly lower dropout hazards than students with ULB grants in years one and three at a 90% confidence level, while students with FRIA doctoral fellowships have significantly lower dropout hazards compared to students with ULB grants in the first four years.

are significantly different at a 90% confidence level. This is in line with the exploratory analysis. The unfinanced students also have higher completion hazards than assistants, which is rather surprising. Four explanations can be given to this phenomenon. First, by taking the decision to pursue a doctorate without financing, these students show their motivation with regard to doctoral study. Second, the high completion hazards of the unfinanced students can be explained with the non-registration problem. Indeed, if students who have no financing tend to register less, it can be that only the unfinanced students who are confident that they will complete their Ph.D. These two explanations are however contradictory to the high dropout hazards of this group of students. Third, it might be that the significantly higher completion coefficient of the unfinanced students is driven by the hazards before year six, which would fit with the fact that assistantships entail teaching duties and are given for *maximum* six years. Finally, it must be kept in mind that the "no financing" category still includes some types of financing, which the administration of the ULB cannot detect. Table 9 gives an overview of the sample hazards by type of financing that was most received until then. From this table we learn that the sample hazard of completing a Ph.D. is indeed only higher for unfinanced students compared to teaching assistants before year six. We can also see that, as said previously, FNRS doctoral fellows and students with FRIA and ULB grants do not necessarily have lower completion rates than teaching assistants after year five. Moreover, the completion hazards of FNRS doctoral fellows and teaching assistants in year four are very different (respectively 34% and 8%) while the hazard of FRIA students in this year is similar to the one of teaching assistants (9%).

INSERT Table 9: Sample hazards by type of financing for the whole sample (N=10,235)

Interestingly, **Result** has no significant influence on completion probabilities, except for a negative impact for the category "without honors or *cum laude*" in year five. This is in contrast with the results of the exploratory analysis and can be explained by the fact that in the graphical analysis, we look at the impact of a variable without controlling for other covariates. When the financing variable is not included in the model, both the *cum laude* and *magna cum laude* categories have significantly lower completion hazards than the *summa cum laude* category in years four and five. Thus, it seems that type of financing received does influence degree attainment, beyond the ability of the student.

The final conclusion from the discussion on type of financing and ability is thus that grants (research fellowship, FRIA and ULB) have higher completion hazards compared to assistantships and no financing. This in turn is partly due to the characteristics of grants. Concerning dropout, grants have lower hazards, although this is not the case for ULB grants, as ability still plays a role.

6.6 Field of study

Looking at the results of the first model (Table 6), the impact of field of study on the hazards of dropout only appears for a few years. This is in contrast with the big difference detected graphically, at least for science and technology. In the second model (Table 7), the influence on dropout is even less clear, as the coefficient of science and technology becomes significantly *higher* than the one of human sciences in year three. The only clear pattern in this model is a lower dropout hazard for science and technology and health sciences in year five[26]. It is interesting to note that this is still the case without the variables **Financing** and **Result** in the model.

Concerning completion, in the first model the hazards are significantly higher for science and technology compared to human sciences from year five onwards, which this time tallies with the exploratory analysis. There is however mostly no significant difference in Ph.D. completion probabilities between health sciences and human sciences. In the second model, the impacts become less significant, with even a significantly *lower* completion hazard for health sciences compared to human sciences in year four. Wald tests of equality of coefficients between health sciences and science and technology in model two indicate significantly higher completion hazards for science and technology in years four, five and six at a 90% confidence level.

In these two models, no interaction terms other than the ones with time dummies were included. However, some variables could have different impacts according to the value taken on by another variable. For example, we saw that type of financing is an important determinant of both dropout and completion hazards. It could be that the impacts of the other variables in the model change according to the value taken on by the financing variable. Therefore, it is interesting to estimate separate models for the different categories of financing. We estimated the second model for five subsamples consisting of person-year lines for which the variable **Financing** took a particular value. As we look at the second model here, there are only person-year lines from the Belgian university subsample. The estimated coefficients of the two most interesting variables in these models, *i.e.* **Female** and **FOS**, are shown in Table 10. First, while gender had no impact in the general model, female FNRS doctoral fellows drop out significantly more and female teaching assistants are significantly less likely to complete their doctorate than their male counterparts. The most interesting results concern the field of study. For FNRS doctoral fellows and teaching assistants, this variable mostly has no impact. Students with ULB grants who pursue a doctorate in health sciences or science and technology even have significantly *higher* hazards of dropout in the first (two) year(s) than their colleagues in human

[26] Wald tests of equality of coefficients between health sciences and science and technology in model two only indicate a significantly *higher* dropout hazard for science and technology in year three at a 90% confidence level.

sciences. The expected lower completion hazards for human sciences compared to science and technology only appears for unfinanced students. Even for the unfinanced students, the field of study has mostly no impact on dropout. Given that the differences in completion between fields of study only exist for unfinanced students, this particular group of students drives the significant results of the two main models in terms of completion hazards for science and technology. Note that even when the second model is estimated for the categories "FNRS doctoral fellowship", "ULB grant" and "assistant" together in order to have more observations, the fields of study categories are still not significant[27]. Moreover, when we include students who did not pursue their second cycle in Belgium in the estimation of the separate models using the variables of the first model, field of study is again insignificant for the estimations for research fellowships, ULB grants and teaching assistants. It is however noteworthy that for FRIA doctoral fellowships, for which the control category of the **FOS** variable is science and technology, health science students have significantly lower completion hazards from year five onwards.

INSERT Table 10: Results of fitting the second model for different types of financing, Belgian university subsample (Selected variables)

In order to have an idea of the marginal effects of the fields of study, the $\Delta\hat{h}(k,t)$ (k=1,2 and t=1,…,9) suggested by Scott and Kennedy (2005) were computed. We chose a particular individual (a male Belgian single student who started a Ph.D. before he turned 26, completed his second cycle at the ULB, obtained a magna cum laude result, did not fail a year and chose a discipline for his Ph.D. that was similar to the one of his second cycle) and estimated, for the "assistant" and "no financing" models, his dropout and completion hazards in each field of study for the nine years. Afterwards, the differences between the estimated hazards for health sciences and science and technology with human sciences were taken. Figure 9 displays those differences over time. We can see that there are no important effects of field of study on dropout hazards, neither for teaching assistants nor for unfinanced students. Concerning completion hazards, there is an important marginal effect of the science and technology category in year six for teaching assistants, but it is not significant at usual confidence levels. However, we can see that unfinanced students in science and technology have completion probabilities which are increased by about 15% in year four, about 45% in year five and almost 50% in year six compared to human science students. The marginal effects of the category health sciences are always close to zero.

[27] Actually, in this estimation the health science students have significantly *higher* dropout hazards than human science students in years one and two and science and technology students in years one and three at a 90% confidence level. Health science students have a significantly *lower* completion hazard in year four compared to human sciences. There are no significant differences between health sciences and science and technology.

INSERT Figure 9: Marginal effects of field of study on hazards for models by type of financing (control category = human sciences)

From the discussion about fields of study three main conclusions can be drawn. First, there are mainly no differences in hazards between human sciences and health sciences. Second, the differences in dropout hazards between fields are small once other control variables are included. Finally, human science and health science students have significantly lower completion hazards in later years, but only in the group of unfinanced students, and of FRIA doctoral fellowships for health sciences.

We can conclude that field of study has no impact when we look at types of financing separately, except for unfinanced students. Thus, at equal levels of ability, human science and health science students should not be more often unfinanced than their science and technology colleagues, which is not the case nowadays. Table 11 displays the percentage of students who were totally unfinanced during their doctoral education by field of study and result obtained in second cycle studies. Human science and health science students are always more often unfinanced than Ph.D. candidates in science and technology. In general the percentage of students without financing is higher in the human sciences compared to health sciences, except for *summa cum laude* students, who have easily access to the FNRS doctoral fellowship. There is thus a need to expand access to financing for doctorates in human sciences and health sciences, which is justified by the fact that receiving some type of financing influences Ph.D. completion beyond the ability of the student. Grants, which have high completion hazards, are important in this respect. A FRIA type of grant could for example be created for human sciences. Moreover, the number of ULB grants for human sciences and health sciences can be increased. FRIA and ULB grants are very important for students who were not awarded the FNRS doctoral fellowship. Indeed, students with these grants have completion hazards that are mostly not significantly different from those of FNRS doctoral fellows, and that are higher than the ones of teaching assistants and unfinanced students. The need to expand financing for human and health sciences is emphasized by the fact that the model for "no financing" was precisely the only one in which field of study had a significant impact on completion. When estimating separate models per field of study, we can actually learn that only for science and technology being unfinanced increases completion hazards.

INSERT Table 11: Percentage of students who were totally unfinanced during their doctoral education by result in second cycle and field of study, Belgian university subsample (n=1,915)

7. Conclusion

Our competing risks survival analysis of the hazards of dropout of doctoral study and Ph.D. completion in each year at the ULB showed that neither gender nor nationality has an impact on the doctoral path. However, it seems that students who finish a second cycle at the ULB take the decision of continuing higher education at the ULB more lightly than others. Thus, it is important to encourage those students to think their decision through and to make sure that they are well informed, in order to avoid allocating financing to students who will drop out. We conclude that FNRS doctoral fellows and students with FRIA doctoral fellowships have the lowest dropout hazards and unfinanced students the highest. However, the influence on dropout hazards is also influenced by the "ability" of the student, and more specifically the result obtained in second cycle studies. Thus, the allocation of financing is efficient in that it is awarded to students who will drop out less. Grants (FNRS, FRIA doctoral fellowships, and ULB grants) have the highest Ph.D. completion hazards, followed by unfinanced students and finally teaching assistants. We showed that receiving a grant is important in terms of Ph.D. completion. However, the influence of ability on Ph.D. completion is more difficult to ascertain. Finally, our results showed that the general idea that human science students are less likely to complete their doctorate and that they take more time if they do so is not grounded. Actually, there are no differences in field of study, except for unfinanced students. It is this group of students who drives the higher dropout and lower completion rates in human sciences. Thus, at equal "ability" – which is difficult to perceive in the context of doctoral study – students in human and health sciences should be given the same access to financing, which should influence their completion probabilities. Grants are important in this respect. The need to expand financing for these fields of study is moreover emphasized by the fact that only unfinanced students in science and technology have higher completion hazards than teaching assistants.

In a context of increasing importance of doctorate holders and of limited resources for research, understanding what determines dropout of doctoral study or completion of the Ph.D. degree is of utmost importance. This article made a contribution in this respect. However, it has some limitations. First, not all variables that could have been worth analyzing were accessible. For example, the number of children, the relationship between thesis supervisor and student, and data about the first cycle of studies could have been interesting variables. Moreover, this study excludes departmental factors, which have been found to be significant by other researchers. If departmental factors, such as number of students, integration of students and availability of faculty, influence dropout or degree completion, it is the duty of departmental committees to improve the efficiency of their doctoral activities given the important role of universities in knowledge societies (Visser *et al.*, 2007). Second, the analysis was restricted to the students of the ULB. However, it could be hypothesized that the associations found here will be similar in other universities of the French Community of Belgium. In any case, this article presented a methodology that can be used in other universities to assess their doctoral process. It would of course be interesting to group all universities of the FCB in one analysis. Finally, only a quantitative study of the doctoral process was performed here. Through interviews with doctoral candidates, doctorate holders, students who dropped out and faculty, a rich understanding of the doctoral process can be gained and the hypothetical explanations presented in this article can be confirmed or invalidated. This is to us the most promising direction of further research. Why do students who studied previously at the ULB drop out more? Why do unfinanced students withdraw more in the first years, but complete their Ph.D. faster than teaching assistants? And why are unfinanced human science and health science students less likely to complete their doctorate than their science and technology colleagues? We tried to put forward answers to these questions, but only thorough interviews with doctoral candidates and faculty will enable us to really understand the patterns detected here and act in consequence. Therefore, our study really is a starting point in the reflection about doctoral study.

Acknowledgements

First of all, we would like to thank ULB authorities for granting the database and especially Mrs. Claude Henschel for the time she spent on the preparation of the database.
Catherine Dehon gratefully acknowledges research support the ARC contract of the Fédération Wallonie-Bruxelles. Catherine Dehon is also member of ECORE, the association between CORE and ECARES.
The article was completed when Elena Arias hold a Short Term Consultant position at the World Bank. The findings and interpretations of the article do not necessarily represent the views of neither the Inter-American Development Bank nor the World Bank. The usual disclaimer applies.

References

Abedi, J. and Benkin, E. (1987), "The Effects of Students' Academic, Financial and Demographic Variables on Time to the Doctorate", *Research in Higher Education*, 27(1), pp. 3-14.

Ai, C. and Norton, E.C. (2003), "Interaction Terms in Logit and Probit Models", *Economics Letters,* 80(1), pp. 123-129.

Allison, P.D. (1982), "Discrete-Time Methods for the Analysis of Event Histories", in Leinhardt S. (Ed.) Sociological Methodology, San Francisco: Jossey-Bass, pp. 61-98.

Arias Ortiz, E. and Dehon, C. (2011), "The Roads to Success: Analyzing Dropout and Degree Completion at University", *mimeo.*

Bair, C.R. and Haworth, J.G. (2005), "Doctoral Student Attrition and Persistence: A Meta-Synthesis of Research", *Higher Education: Handbook of Theory and Research*, 19, pp. 481-534.

Baird, L.L. (1990), "Disciplines and Doctorates: The Relationships between Program Characteristics and the Duration of Doctoral Study", *Research in Higher Education*, 31(4), pp. 369-385.

Booth, A.L. and Satchell, S.E. (1995), "The Hazards of Doing a Ph.D.: An Analysis of Completion and Withdrawal Rates of British Ph.D. Students in the 1980s*", Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158(2), pp. 297-318.

Buis, M.L. (2010), "Stata Tip 87: Interpretation of Interactions in Non-Linear Models*", The Stata Journal*, 10(2), pp. 305-308.

Cox, D.R. (1972), "Regression Models and Life-Tables", *Journal of the Royal Statistical Society*, Series B (Methodological), 34(2), pp. 187-220.

De Henau, J. and Meulders, D. (2003), *Alma mater, homo sapiens? Quel genre pour la recherche universitaire?*, Brussels Economic Series, (Ed.)

Durez, M., Verheve, D. and Hondekyn, I. (2001), "De la thèse à l'emploi en Communauté française de Belgique" in Dewatripont, M., Thys-Clément, F. and Wilkin, L. (Eds.) The Strategic Analysis of Universities: Microeconomic and Managamenent Perspectives, Editions de l'Université de Bruxelles, Brussels, pp. 297-316.

Ehrenberg, R.G. and Mavros, P.G. (1995), "Do Doctoral Students' Financial Support Patterns affect their Times-to-Degree and Completion Probabilities?" *The Journal of Human Resources*, 30(3), pp. 581-609.

Eurostat (2011), *Science, Technology and Innovation in Europe*, Publications Office of the European Union, Luxembourg.

Ferrer de Valero, Y. (2001), "Departmental Factors affecting Time-to-Degree and Completion Rates of Doctoral Students at one Land-Grant Research Institution", *The Journal of Higher Education*, 72(3), pp. 341-367.

Flinn, C.J. and Heckman, J.J. (1982), "Models for the Analysis of Labor Force Dynamics", NBER Working Paper Series, Working Paper No. 857.

Gillingham, L., Seneca, J.J. and Taussig, M.K. (1991), "The Determinants of Progress to the Doctoral Degree", *Research in Higher Education*, 32(4), pp. 449-468.

Golde, C.M. (2005), "The Role of the Department and Discipline in Doctoral Student Attrition: Lessons from Fours Departments", *The Journal of Higher Education*, 76(6), pp. 669-700.

Grove, W.A., Dutkowsky, D.H. and Grodner, A. (2007), "Survive then Thrive: Determinants of Success in the Economics Ph.D. Program", *Economic Inquiry*, 45(4), pp. 864-871.

Hoetker, G. (2007), "The Use of Logit and Probit Models in Strategic Management Research: Critical Issues", *Strategic Management Journal*, 28(4), pp. 331-343.

Lunneborg, C.E. and Lunneborg, P.W. (1973), "Doctoral Study Attrition in Psychology", *Research in Higher Education*, 1(4), pp. 379-387.

Meulders, D., Plasman, R. and Rycx, F. (2005), "Les inégalités salariales de genre: expliquer l'injustifiable ou justifier l'inexplicable", *Reflets et Perspectives de la vie économique*, 2, pp. 95-107.

Mooney, J.D. (1968), "Attrition among Ph.D. Candidates: An Analysis of a Cohort of Recent Woodrow Wilson Fellows", *The Journal of Human Resources*, 3(1), pp. 47-62.

Moortgat, P. (2011), "CDH (Careers of Doctorate Holders)", Belgian Science Policy Office, D/2011/1191/2.

Nelson, C. and Lovitts, B.E. (2001), "10 Ways to Keep Graduate Students from Quitting", *The Chronicle of Higher Education*, 29 June 2001, p. B20.

Nerad, M. and Cerny, J. (1993), "From Facts to Action: Expanding the Graduate Division's Educational Role", *New Directions for Institutional Research*, 80, pp. 27-39.

Norton, E.C., Wang, H. and Ai, C. (2004), "Computing Interaction Effects and Standard Errors in Logit and Probit Models", *The Stata Journal*, 4(2), pp. 154-167.

O'Dorchai, S.P. (2009), "Do Women Gain or Lose from Becoming Mothers? A Comparative Wage Analysis in 20 European Countries", Working Paper DULBEA, Research Report, N°09-11.RS.

Ott, M.D., Markewich, T.S. and Ochsner, N.L. (1984), "Logit Analysis of Graduate Student Retention", *Research in Higher Education*, 21(4), pp. 439-460.

Park, C. (2005), "War of Attrition: Patterns of Non-Completion amongst Postgraduate Research Students", *Higher Education Review*, 38(1), pp. 48-53.

Scott, M.A. and Kennedy, B.B. (2005), "Pitfalls in Pathways: Some Perspectives on Competing Risks Event History Analysis in Education Research", *Journal of Educational and Behavioral Statistics*, 30(4), pp. 413-442.

Seagram, B.C., Gould, J. and Pyke, S.W. (1998), "An Investigation of Gender and Other Variables on Time to Completion of Doctoral Degrees", *Research in Higher Education*, 39(3), pp. 319-335.

Siegfried, J.J. and Stock, W.A. (2001), "So You Want to Earn a Ph.D. in Economics? How Long Do You Think it Will Take?", *The Journal of Human Resources*, 36(2), pp. 364-378.

Singer, J.D. and Willett, J.B. (1993), "It's about Time: Using Discrete-Time Survival Analysis to study Duration and the Timing of Events", *Journal of Educational Statistics*, 18(2), pp. 155-195.

Smallwood, S. (2004), "Doctor Dropout", *Chronicle of Higher Education*, 50(19), pp. A10-A12.

Stock, W.A., Finegan, T.A. and Siegfried, J.J. (2006), "Attrition in Economics Ph.D. programs", *The American Economic Review*, 96(2), pp. 458-466.

Stricker, L.J. (1994), "Institutional Factors in Time to the Doctorate", *Research in Higher Education*, 35(5), pp. 569-587.

Train, K. (1986), *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*, MIT Press: Cambridge, MA.

Turnbull, B.W. (1974), "Nonparametric Estimation of a Survivorship Function with Doubly Censored Data", *Journal of the American Statistical Association*, 69, pp. 169-173.

Van Ours, J.C. and Ridder, G. (2003), "Fast Track or Failure: A Study of the Graduation and Dropout Rates of Ph D Students in Economics", *Economics of Education Review*, 22(2), pp. 157-166.

Visser, M.S., Luwel, M. and Moed, H.F. (2007), "The Attainment of Doctoral Degrees at Flemish Universities: A Survival Analysis", *Higher Education*, 54(5), pp. 741-757.

Willett, J.B. and Singer, J.D. (1991), "From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition", *Review of Educational Research*, 61(4), pp. 407-450.

Wilson, K.M. (1965), "Of Time and the Doctorate: Report of an Inquiry into the Duration of Doctoral Study", SREB Research Monograph 9, Atlanta: Southern Regional Education Board.

Figures

FIGURE 1: *Maximum likelihood estimation of the hazard functions (whole sample)*

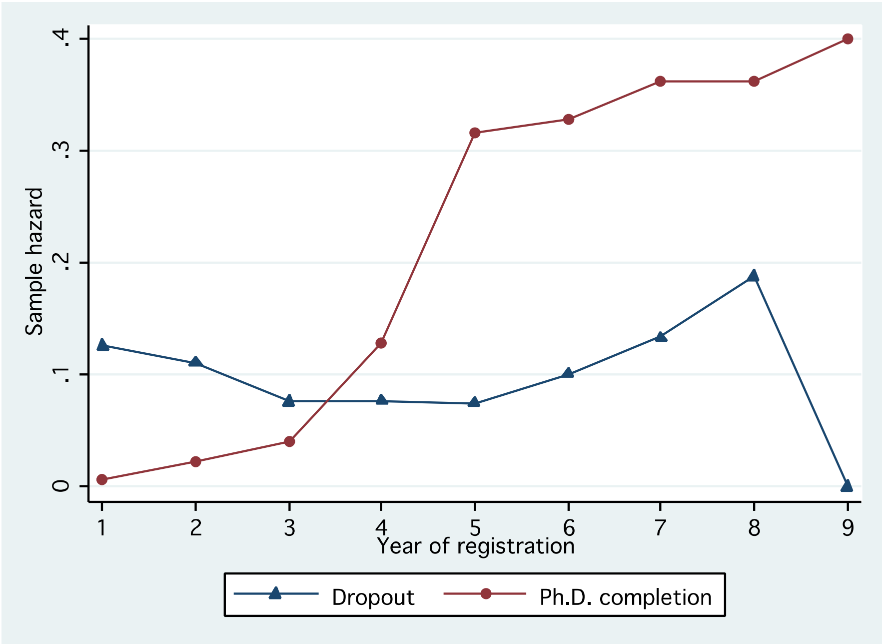


FIGURE 2: *Maximum likelihood estimation of the cumulative hazard functions (whole sample)*

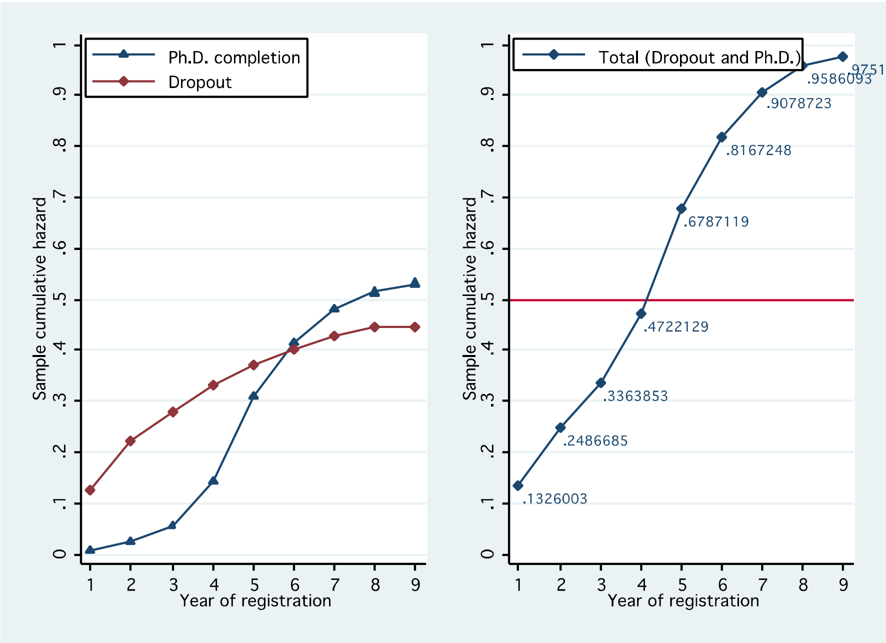


FIGURE 3: *Hazard functions by gender (whole sample)*

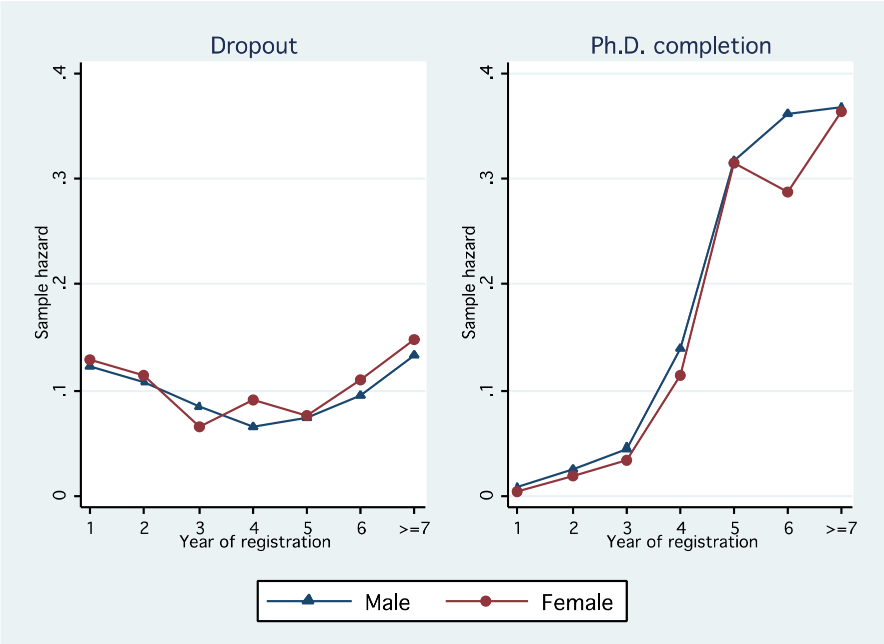


FIGURE 4: *Cumulative hazard functions by field of study (whole sample)*

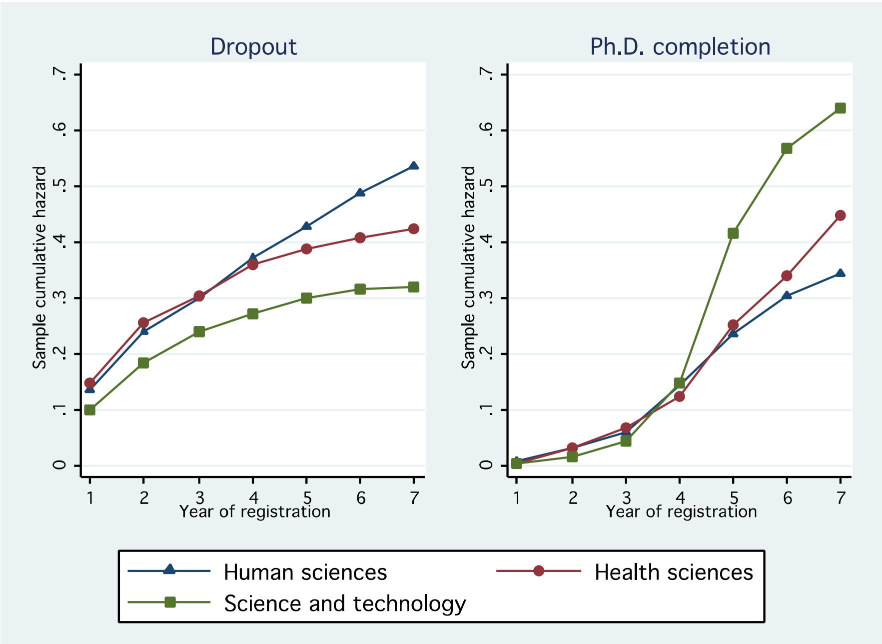


FIGURE 5: *Cumulative hazard functions by type of financing (whole sample)*

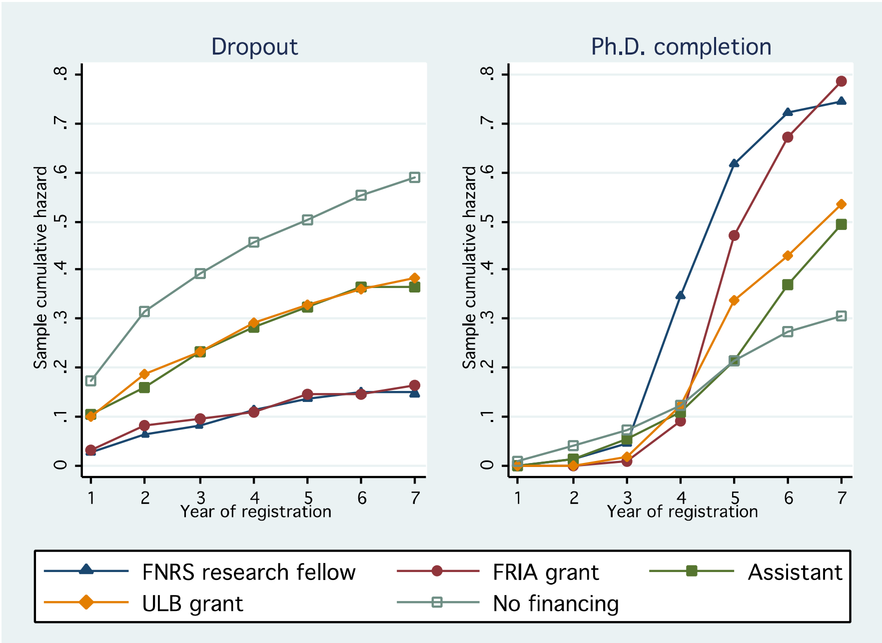


FIGURE 6: Hazard functions by age at start of Ph.D. (whole sample, 26 is the median age at start of Ph.D. in the sample).

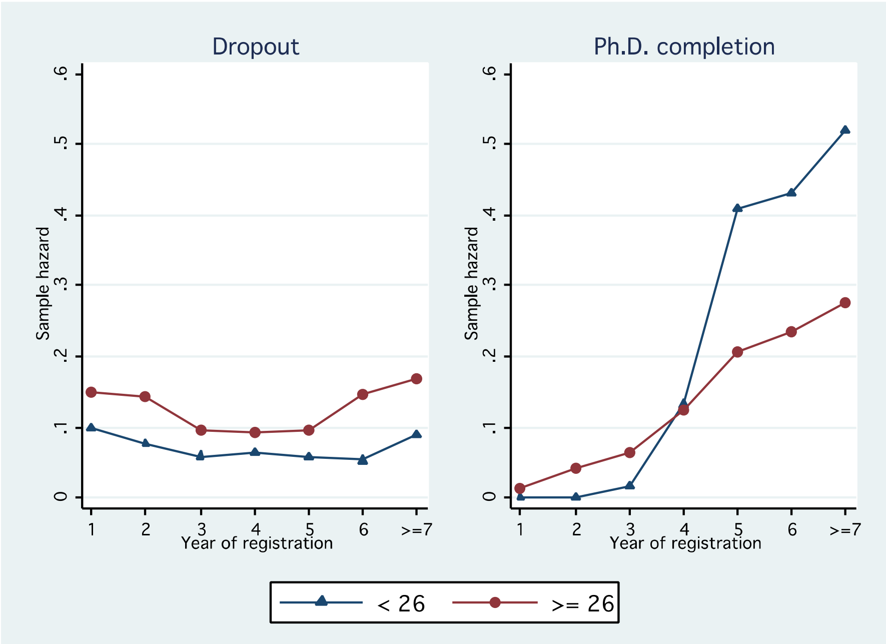


FIGURE 7: *Cumulative hazard functions by result obtained in second cycle studies (Belgian university subsample)*

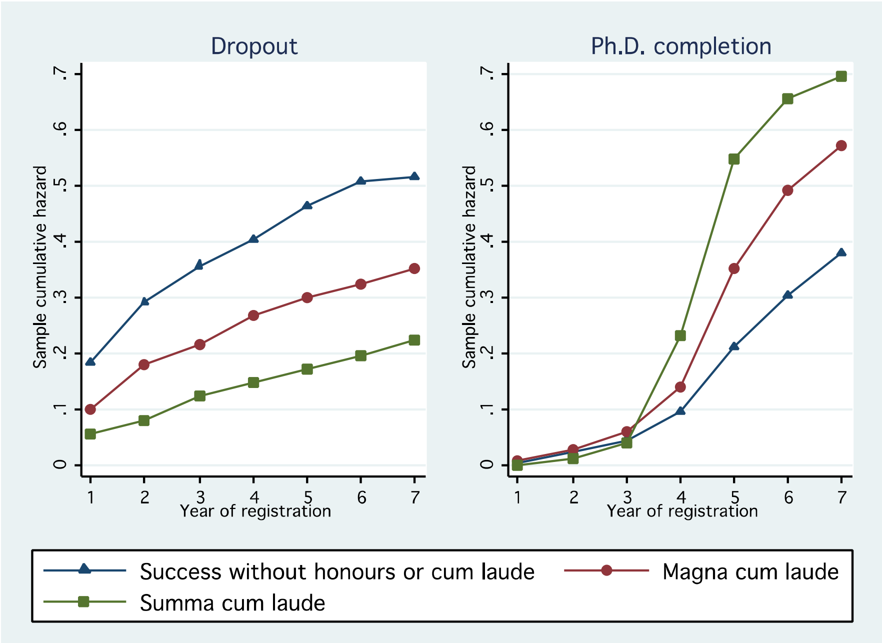
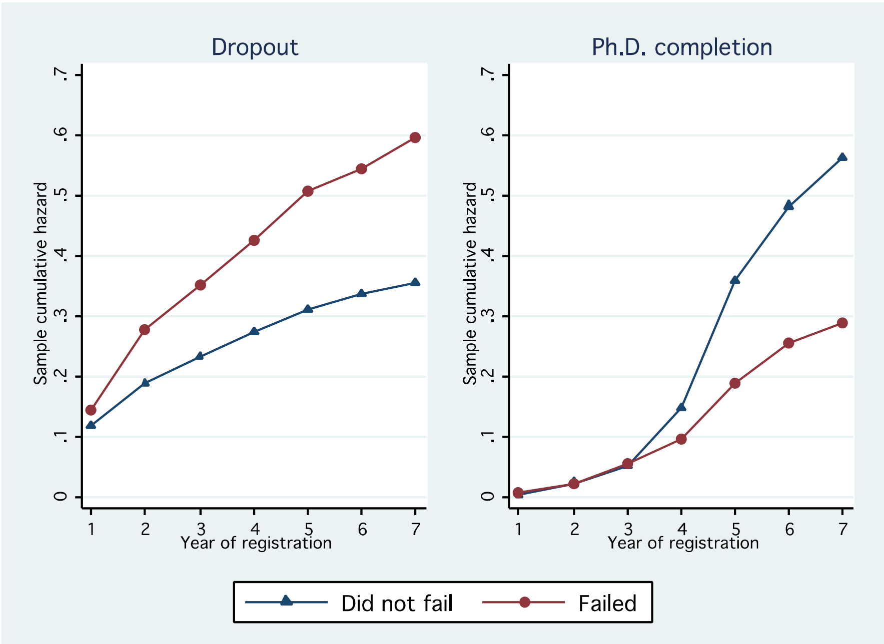


FIGURE 8: *Cumulative hazard functions by failure in second cycle (Belgian university subsample)*



Tables

TABLE 1: *Presentation of the most common types of financing for the pursuit of a doctorate in the French Community of Belgium*

Research fellowship from the FNRS	
Description	The FRS-FNRS, funded mainly by the government of the FCB and the federal government, finances researcher contracts in universities of the FCB. As the first step in the career of a researcher is the completion of a doctorate, the fund awards grants for doctoral study in all disciplines. An FNRS research fellow is not allowed to accept teaching duties.
Length of time	The grant is awarded for two years and is renewable for *maximum* two years. The aim is thus to complete the doctorate in four years.
Selectivity	The allocation of the positions is highly competitive. The selection occurs between different universities and the results obtained in first and second cycle studies are important determinants for the allocation.
FRIA grant from the FNRS	
Description	The FNRS also awards "FRIA" (*Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture*) grants for students who pursue a doctoral degree in science and technology or health sciences. A student benefiting from the grant cannot accept teaching duties.
Length of time	The grant is awarded for 27 months and is renewable for *maximum* 21 months. The aim is thus to complete the doctorate in four years.
Selectivity	The allocation of the grants is made between different universities and is therefore competitive, but less than for FNRS research fellowships. The student must enter an application file and take an oral exam. Moreover, the opinion of a professor and of the thesis supervisor of the student are taken into account. However, the results obtained in first and second cycle studies are still important determinants for the allocation. As the amount awarded by an FNRS research fellowship is slightly higher, students often first apply for this type of grant before applying for the FRIA grant.
Grant from the university	
Description	The university in which the Ph.D. candidate will pursue his doctoral education can also allocate grants. At the ULB, the funds for these grants mainly come from donations to the ULB or from research contracts between the ULB and a firm or government. A student benefiting from an ULB grant cannot accept teaching duties. ULB grants exist in all disciplines.
Length of time	In general, the grant is awarded for two years and is renewable for *maximum* two years. The aim is thus to complete the doctorate in four years.
Selectivity	The allocation of the university grants is less competitive than for FNRS research fellowships and FRIA grants, as it only concerns one university. Moreover, often students first try to obtain a grant from the FNRS. In the case of a research contract, it is in general the thesis supervisor who decides to award the grant to a student, with the approval of the department. In the case of donations, a selection committee decides of the allocation of the grant.
Assistantship	
Description	Some students have an assistantship position paid by the university out of its operating budget. Assistantships can be part-time (0.5 FTE) or full-time (1 FTE) and exist in all disciplines. An assistant dedicates at least 50% of his time to research. The other 50% consists mainly of teaching duties.
Length of time	An assistantship position lasts for two years. It can be renewed *maximum* twice for a full-time position and five times for a part-time position (without exceeding six years in FTE). The aim is thus to complete the doctorate in six years.
Selectivity	The allocation of the assistantship positions is less competitive than for FNRS research fellowships or FRIA grants, as it only concerns one university. Moreover, often students first try to obtain a grant from the FNRS, because they consider them to be more interesting.

TABLE 2: *Descriptive statistics for the whole sample based on number of individuals (n = 3,092)*

Variable	Proportion of individuals in sample	Final outcome in 2009 subsample %		
		Dropout	Ph.D. completion	Censored
Total	100.00%	33.76%	27.17%	39.07%
Gender				
Male	55.37%	32.89%	28.56%	38.55%
Female	44.63%	34.86%	25.43%	39.71%
Nationality				
Belgian	58.28%	30.69%	29.80%	39.51%
Foreigner	41.72%	38.06%	23.49%	38.45%
University of second cycle				
ULB	54.88%	31.29%	29.11%	39.60%
Other university	45.12%	36.77%	24.80%	38.42%
Financing				
FNRS research fellow	8.76%	10.33%	38.01%	51.66%
FRIA grant	12.74%	12.44%	42.39%	45.18%
Assistant	11.16%	27.25%	24.93%	47.83%
ULB grant	14.46%	30.65%	29.31%	40.04%
Researcher	5.24%	35.19%	35.19%	29.63%
Other type of financing	3.62%	32.14%	32.14%	35.71%
No financing	44.02%	47.24%	19.10%	33.65%
Field of study				
Human sciences	40.56%	39.71%	20.18%	40.11%
Health sciences	21.25%	35.62%	26.94%	37.44%
Science and technology	38.20%	26.42%	34.72%	38.87%
Age				
Start of Ph.D. prior to age 26	46.18%	26.26%	31.30%	42.44%
Start of Ph.D. at age 26+	53.82%	40.20%	23.62%	36.18%
Marital status				
Not married	85.21%	32.52%	27.48%	40.00%
Married	14.78%	40.92%	25.38%	33.70%
Year of start of Ph.D.				
2001-2002	10.22%	46.84%	50.32%	2.85%
2002-2003	11.38%	41.19%	54.26%	4.55%
2003-2004	11.64%	45.56%	47.50%	6.94%
2004-2005	9.67%	37.46%	42.47%	20.07%
2005-2006	10.54%	36.20%	34.36%	29.45%
2006-2007	12.71%	29.77%	14.50%	55.73%
2007-2008	12.68%	32.65%	3.32%	64.03%
2008-2009	9.96%	25.65%	3.25%	71.10%
2009-2010	11.19%	9.54%	0.00%	90.46%

Notes:

26 is the median age at start of Ph.D. in the sample.

Financing and marital status are time-varying variables, which means that they can take on different values over time for the same individual. For these descriptive statistics, one value was selected per individual. For type of financing, the value selected was the type of financing that was most received during the period of doctoral study. If a student received two or more types of financing during an equal number of years, one of these types of financing was selected based on the following order: FNRS research fellow, FRIA grant, ULB grant, assistant, researcher, other type of financing and no financing. For marital status, the value selected was the one that the individual had in his last year of registration.

TABLE 3: *Descriptive statistics for the Belgian university subsample based on number of individuals, only "second cycle variables" (n = 1,915)*

Variable	Proportion of individuals in subsample	Final outcome in 2009 subsample %		
		Dropout	Ph.D. completion	Censored
Total	100.00%	30.60%	29.40%	40.00%
Result obtained in second cycle studies				
Success without honours	4.18%	48.75%	18.75%	32.50%
Cum laude	29.14%	41.40%	22.04%	36.56%
Magna cum laude	50.23%	27.34%	31.60%	41.06%
Summa cum laude	16.45%	16.83%	38.41%	44.76%
Equivalence between discipline of second cycle and of Ph.D.				
Equivalence	91.07%	29.99%	29.99%	40.02%
No equivalence	8.93%	36.84%	23.39%	39.77%
Failed (a) year(s) during second cycle studies				
Did not fail a year	87.26%	28.31%	31.12%	40.57%
Failed at least one year	12.74%	46.31%	17.62%	36.07%

TABLE 4: *Presentation of survival data for the whole sample (n = 3,092)*

	Number who				Proportion of	
Year	Were still in the doctoral process at the beginning of the year ("risk set")	Dropped out during the year	Completed their Ph.D. during the year	Were censored at the end of the year	Students at the beginning of the year who dropped out by the end of the year (MLE for the hazard probability of dropout in that year)	Students at the beginning of the year who completed their Ph.D. by the end of the year (MLE for the hazard probability of Ph.D. completion in that year)
1	3,092	389	21	313	0.126	0.007
2	2,369	263	54	219	0.111	0.023
3	1,833	140	74	251	0.076	0.040
4	1,368	105	175	219	0.077	0.128
5	869	65	275	96	0.075	0.316
6	433	44	142	60	0.102	0.328
7	187	25	68	25	0.134	0.364
8	69	13	25	16	0.188	0.362
9	15	0	6	9	0.000	0.400

TABLE 5: *Results of fitting the baseline model using the whole sample*

	Year of registration for doctoral study						
	1	2	3	4	5	6	≥7
Coefficients							
Dropout	-1.931	-2.054	-2.448	-2.338	-2.097	-1.725	-1.260
Ph.D. completion	-4.850	-3.638	-3.085	-1.827	-0.654	-0.554	-0.303
Outcome-specific hazard ratios							
Dropout	0.145	0.128	0.086	0.097	0.123	0.178	0.284
Ph.D. completion	0.008	0.026	0.046	0.161	0.520	0.575	0.739
Outcome-specific hazards							
Dropout	0.126	0.111	0.076	0.077	0.075	0.102	0.140
Ph.D. completion	0.007	0.023	0.040	0.128	0.317	0.328	0.365
N				10,235			

Note: All coefficients are significant at a 99% confidence level, except years ≥7 for the Ph.D. outcome, which is significant at a 95% confidence level.

TABLE 6: *Results of fitting the first model using the whole sample*

	Dropout		Ph.D. completion	
	Coefficient	Exponentiated coefficient	Coefficient	Exponentiated coefficient
Female	0.083	1.087	-0.041	0.960
Belgian	-0.017	0.983	0.125	1.133
ULB	0.179 **	1.196	0.040	1.041
Aged * year 1	0.153	1.165	/	/
Aged * year 2	0.243	1.275	/	/
Aged * year 3	-0.117	0.890	1.308 ***	3.699
Aged * year 4	-0.320	0.726	0.245	1.278
Aged * year 5	0.101	1.106	-0.390 *	0.677
Aged * year 6	0.450	1.568	-0.121	0.886
Aged * years 7 8 9	-0.247	0.781	-0.733 **	0.480
Health sciences * year 1	0.279 **	1.322	-0.411	0.663
Health sciences * year 2	0.187	1.206	0.268	1.307
Health sciences * year 3	0.012	1.012	0.319	1.376
Health sciences * year 4	-0.080	0.923	-0.414	0.661
Health sciences * year 5	-0.833 **	0.435	0.164	1.178
Health sciences * year 6	-0.843 *	0.430	-0.058	0.944
Health sciences * years 7 8 9	-0.696	0.499	0.760 **	2.138
Science and technology * year 1	0.036	1.037	-0.423	0.655
Science and technology * year 2	0.098	1.103	0.143	1.154
Science and technology * year 3	0.328	1.388	0.489 *	1.631
Science and technology * year 4	-0.588 **	0.555	0.334	1.397
Science and technology * year 5	-0.436	0.647	1.102 ***	3.010
Science and technology * year 6	0.003	1.003	1.216 ***	3.374
Science and technology * years 7 8 9	-1.401 *	0.246	1.073 **	2.924
MSd * year 1	-0.427 ***	0.652	-0.888	0.411
MSd * year 2	0.155	1.168	0.179	1.196
MSd * year 3	0.152	1.164	-0.254	0.776
MSd * year 4	0.268	1.307	0.517 **	1.677
MSd * year 5	-0.666	0.514	0.268	1.307
MSd * year 6	-0.375	0.687	-0.752 *	0.471
MSd * years 7 8 9	0.376	1.456	0.413	1.511

TABLE 6 (continued): *Results of fitting the first model using the whole sample*

	Dropout			Ph.D. completion		
	Coefficient		Exponentiated coefficient	Coefficient		Exponentiated coefficient
FNRS research fellow * year 1	-1.326	***	0.266	/		/
FNRS research fellow * year 2	-0.872	**	0.418	0.330		1.391
FNRS research fellow * year 3	-1.309	**	0.270	0.453		1.573
FNRS research fellow * year 4	-0.416		0.660	2.109	***	8.240
FNRS research fellow * year 5	0.084		1.088	1.643	***	5.171
FNRS research fellow * year 6	-0.165		0.848	0.866	*	2.377
FNRS research fellow * years 7 8 9	-0.391		0.676	-0.487		0.614
FRIA grant * year 1	-1.125	***	0.325	/		/
FRIA grant * year 2	-0.495	*	0.610	-0.488		0.614
FRIA grant * year 3	-2.005	***	0.135	-0.918		0.399
FRIA grant * year 4	-1.080	**	0.340	0.283		1.327
FRIA grant * year 5	0.419		1.520	0.928	***	2.529
FRIA grant * year 6	/		/	0.626	*	1.870
FRIA grant * years 7 8 9	1.113		3.043	1.046	**	2.846
ULB grant * year 1	0.038		1.039	/		/
ULB grant * year 2	0.116		1.123	/		/
ULB grant * year 3	-0.282		0.754	-0.380		0.684
ULB grant * year 4	0.375		1.455	0.884	***	2.421
ULB grant * year 5	0.469		1.598	1.027	***	2.793
ULB grant * year 6	0.320		1.377	0.139		1.149
ULB grant * years 7 8 9	0.571		1.770	1.067	**	2.907
Researcher	0.303	*	1.354	0.994	***	2.702
Other financing	0.260		1.297	0.668	***	1.950
No financing	0.756	***	2.130	0.409	**	1.505
Time dummies	Yes			Yes		
N	10,235					

***: The coefficient is significantly different from 0 at a 99% confidence level

** : The coefficient is significantly different from 0 at a 95% confidence level

* : The coefficient is significantly different from 0 at a 90% confidence level

Note: "/" means that the estimate is based on too few observations. For example, one assistant obtained his doctorate in the first year, while no FNRS research fellow did.

TABLE 7: *Results of fitting the second model using the Belgian university subsample (Selected variables)*

	Dropout			Ph.D. completion		
	Coefficient		Exponentia-ted coefficient	Coefficient		Exponentia-ted coefficient
Wihout honours or *cum laud*e * year 1	0.710	**	2.034	-0.317		0.728
Wihout honours or *cum laude* * year 2	1.283	***	3.607	-0.420		0.657
Wihout honours or *cum laude* * year 3	0.077		1.080	-0.544		0.580
Wihout honours or *cum laude* * year 4	0.646		1.908	-0.157		0.855
Wihout honours or *cum laude* * year 5	0.839		2.314	-0.633	**	0.531
Wihout honours or *cum laude* * year 6	0.452		1.571	-0.082		0.921
Wihout honours or *cum laude* * years 7 8 9	-1.835	**	0.160	0.217		1.242
Magna cum laude * year 1	0.260		1.297	0.419		1.520
Magna cum laude * year 2	1.031	**	2.804	0.023		1.023
Magna cum laude * year 3	-0.443		0.642	-0.086		0.918
Magna cum laude * year 4	0.835		2.305	0.049		1.050
Magna cum laude * year 5	0.138		1.148	-0.336		0.715
Magna cum laude * year 6	-0.074		0.929	0.014		1.014
Magna cum laude * years 7 8 9	-0.689		0.502	0.311		1.365
Fail	0.296	**	1.344	-0.405	**	0.667
Equivalence	-0.112		0.894	-0.175		0.839
FNRS research fellow * year 1	-1.144	**	0.319	/		/
FNRS research fellow * year 2	-0.396		0.673	0.349		1.418
FNRS research fellow * year 3	-1.381	**	0.251	0.400		1.492
FNRS research fellow * year 4	0.120		1.127	2.427	***	11.325
FNRS research fellow * year 5	0.147		1.158	1.742	***	5.709
FNRS research fellow * year 6	0.182		1.200	0.858		2.358
FNRS research fellow * years 7 8 9	-0.405		0.667	-0.575		0.563
FRIA grant * year 1	-1.297	***	0.273	/		/
FRIA grant * year 2	-0.398		0.672	-0.196		0.822
FRIA grant * year 3	-2.363	***	0.094	-0.546		0.579
FRIA grant * year 4	-1.711	***	0.181	0.601		1.824
FRIA grant * year 5	1.018	*	2.768	1.309	***	3.702
FRIA grant * year 6	/		/	0.799	**	2.223
FRIA grant * years 7 8 9	1.286		3.618	1.055	*	2.872
ULB grant * year 1	-0.202		0.817	/		/
ULB grant * year 2	0.259		1.296	/		/
ULB grant * year 3	-0.076		0.927	-1.677		0.187
ULB grant * year 4	0.511		1.667	0.861	**	2.366
ULB grant * year 5	0.710		2.034	1.580	***	4.855
ULB grant * year 6	0.260		1.297	0.079		1.082
ULB grant * years 7 8 9	1.214		3.367	0.838		2.312
Researcher	0.153		1.165	1.060	***	2.886
Other financing	0.429	*	1.536	0.802	***	2.230
No financing	0.665	***	1.944	0.400	**	1.492

TABLE 7 (continued): *Results of fitting the second model using the Belgian university subsample (Selected variables)*

	Dropout			Ph.D. completion		
	Coefficient		Exponentia-ted coefficient	Coefficient		Exponentia-ted coefficient
Health sciences * year 1	0.254		1.289	-0.157		0.855
Health sciences * year 2	0.285		1.330	0.493		1.637
Health sciences * year 3	-0.048		0.953	0.038		1.039
Health sciences * year 4	0.041		1.042	-0.783	**	0.457
Health sciences * year 5	-1.361	***	0.256	0.034		1.035
Health sciences * year 6	-0.531		0.588	-0.363		0.696
Health sciences * years 7 8 9	-0.648		0.523	0.733		2.081
Science and technology * year 1	0.145		1.156	/		/
Science and technology * year 2	0.322		1.380	0.945	*	2.573
Science and technology * year 3	0.620	**	1.859	0.208		1.231
Science and technology * year 4	-0.071		0.931	0.142		1.153
Science and technology * year 5	-0.852	*	0.427	0.773	***	2.166
Science and technology * year 6	-0.656		0.519	0.658	*	1.931
Science and technology * years 7 8 9	-1.158		0.314	1.109	**	3.031
Time dummies	Yes			Yes		
N	6,492					

***: The coefficient is significantly different from 0 at a 99% confidence level

** : The coefficient is significantly different from 0 at a 95% confidence level

* : The coefficient is significantly different from 0 at a 90% confidence level

Note: "/" means that the estimate is based on too few observations. For example, one assistant obtained his doctorate in the first year, while no FNRS research fellow did.

TABLE 8: *Contingency table between **Financing** and **Result** based on number of individuals for the Belgian university subsample (n = 1,915)*

Financing / Result	No honours	*Cum laude*	*Magna cum laude*	*Summa cum laude*	Total
FNRS research fellow	3	7	92	153	255
Percent (column)	*3.75*	*1.25*	*9.56*	*48.57*	*13.32*
Percent (row)	*1.18*	*2.75*	*36.08*	*60.00*	*100.00*
FRIA grant	3	73	216	57	349
Percent (column)	*3.75*	*13.08*	*22.45*	*18.10*	*18.22*
Percent (row)	*0.86*	*20.92*	*61.89*	*16.33*	*100.00*
Assistant	7	94	171	27	299
Percent (column)	*8.75*	*16.85*	*17.78*	*8.57*	*15.61*
Percent (row)	*2.34*	*31.44*	*57.19*	*9.03*	*100.00*
ULB grant	6	78	143	26	253
Percent (column)	*7.50*	*13.98*	*14.86*	*8.25*	*13.21*
Percent (row)	*2.37*	*30.83*	*56.52*	*10.28*	*100.00*
Researcher	6	36	66	9	117
Percent (column)	*7.50*	*6.45*	*6.86*	*2.86*	*6.11*
Percent (row)	*5.13*	*30.77*	*56.41*	*7.69*	*100.00*
Other financing	3	22	48	10	83
Percent (column)	*3.75*	*3.94*	*4.99*	*3.17*	*4.33*
Percent (row)	*3.61*	*26.51*	*57.83*	*12.05*	*100.00*
No financing	52	248	226	33	559
Percent (column)	*65.00*	*44.44*	*23.49*	*10.48*	*29.19*
Percent (row)	*9.30*	*44.36*	*40.43*	*5.90*	*100.00*
Total	80	558	962	315	1,915
Percent (column)	*100.00*	*100.00*	*100.00*	*100.00*	*100.00*
Percent (row)	*4.18*	*29.14*	*50.23*	*16.45*	*100.00*

Note: The **Financing** category assigned to an individual is the type of financing that he received most during his whole doctoral education.

TABLE 9: *Sample hazards by type of financing for the whole sample (N = 10,235)*

	FNRS research fellow		FRIA grant		ULB grant		Assistant		No financing	
Year	Drop.	Ph.D.	Drop.	Ph.D.	Drop.	Ph.D.	Drop.	Ph.D.	Drop.	Ph.D.
1	2.88%	0.00%	3.40%	0.00%	10.16%	0.00%	10.79%	0.32%	17.27%	0.88%
2	3.57%	1.34%	5.11%	0.28%	9.54%	0.00%	5.86%	1.56%	17.27%	3.84%
3	2.14%	3.74%	1.29%	0.97%	5.52%	2.41%	9.05%	4.76%	12.21%	5.00%
4	3.65%	34.31%	1.98%	9.13%	8.15%	13.73%	6.75%	7.98%	11.75%	9.62%
5	4.29%	50.00%	4.52%	47.46%	6.47%	36.69%	7.14%	16.96%	10.96%	21.59%
6	4.76%	42.86%	0.00%	52.78%	9.59%	27.40%	9.09%	33.33%	17.28%	20.37%
7-9	9.10%	18.18%	7.14%	64.29%	10.00%	52.50%	2.38%	40.48%	21.37%	25.95%

Note: The **Financing** category assigned to an individual is the type of financing that he received most during his whole doctoral education.

TABLE 10: *Results of fitting the second model for different types of financing, Belgian university subsample (Selected variables)*

	FNRS research fellow	FRIA grant (control category **FOS** = S&T)	ULB grant	Assistant	No financing
	Dropout coefficients				
Female	0.969 *	0.308	0.254	0.144	0.131
Health sciences * year 1	0.624	0.769	2.377 ***	1.039 *	-0.352
Health sciences * year 2	0.004	-1.565	1.298 **	0.086	0.069
Health sciences * year 3	/	/	1.327	0.402	-0.400
Health sciences * year 4	-0.569	0.707	1.147	-0.756	0.165
Health sciences * year 5	/	0.611	/	/	-1.668 **
Health sciences * year 6	/	-0.597	-0.979	1.648	-1.324
Health sciences * years 7 8 9	/	-2.425	/	/	-1.261
Science and technology * year 1	0.299		2.444 ***	0.825 *	-0.403
Science and technology * year 2	0.958		0.897	-0.494	0.333
Science and technology * year 3	0.406		1.060	1.000	0.280
Science and technology * year 4	/		-0.104	-0.387	0.013
Science and technology * year 5	0.191		-0.224	0.236	-1.336
Science and technology * year 6	-2.616		/	-0.665	0.575
Science and technology * years 7 8 9			/	/	/
	Ph.D. completion coefficients				
Female	-0.389	-0.038	-0.428	-1.006 ***	0.372
Health sciences * year 1	0.190	0.052	0.055	/	0.270
Health sciences * year 2	0.502	/	-0.074	/	0.921
Health sciences * year 3	0.582	/	/	-0.253	0.564
Health sciences * year 4	-1.365 *	-1.717	-1.301	1.064	0.751
Health sciences * year 5	-0.264	-1.464 ***	0.569	-0.587	1.065 *
Health sciences * year 6	1.683	-1.404 **	-0.835	-0.096	-0.428
Health sciences * years 7 8 9	/	-2.578 *	/	/	0.091
Science and technology * year 1	-0.269		-0.045	/	/
Science and technology * year 2	2.152		-0.420	1.125	1.432 *
Science and technology * year 3	-0.795		/	-1.041	1.475 *
Science and technology * year 4	-0.136		-0.206	-1.650	1.614 **
Science and technology * year 5	-0.291		-0.112	-0.285	2.628 ***
Science and technology * year 6	/		-0.632	1.115	2.247 **
Science and technology * years 7 8 9			/	0.536	-0.271
N	799	1,387	894	1,024	1,707

***: The coefficient is significantly different from 0 at a 99% confidence level

** : The coefficient is significantly different from 0 at a 95% confidence level

* : The coefficient is significantly different from 0 at a 90% confidence level

Note: "/" means that the estimate is based on too few observations.

TABLE 11: *Percentage of students who were totally unfinanced during their doctoral education by result in second cycle and field of study, Belgian university subsample (n = 1,915)*

Result / Field of study	Human sciences	Health sciences	Science and technology	All
Summa cum laude	10.85%	16.98%	4.51%	9.21%
Magna cum laude	31.44%	26.57%	9.95%	21.41%
Cum laude	57.26%	36.36%	21.91%	41.04%
No honours	90.00%	50.00%	22.73%	62.50%
All	39.35%	29.51%	12.24%	26.84%

TABLE 12: *Presentation of fields of study and disciplines, whole sample (n = 3,092)*

Field of study / Discipline	Number of individuals in sample	Proportion of individuals in sample
Human sciences	1,254	40.56%
Philosophy	104	3.36%
Languages and letters	131	4.24%
History, art and archaeology	188	6.08%
Information and communication	64	2.07%
Political and social sciences	361	11.68%
Juridic sciences	45	1.46%
Criminology	11	0.36%
Economic and management sciences	176	5.69%
Psychological and educational sciences	173	5.60%
Art and art sciences	1	0.03%
Health sciences	657	21.25%
Dental sciences	15	0.49%
Mobility sciences	28	0.91%
Medical sciences (Medicine department)	218	7.05%
Medical sciences (Public health department)	59	1.91%
Biomedical and pharmaceutical sciences (Medicine department)	293	9.48%
Biomedical and pharmaceutical sciences (Pharmacy department	44	1.42%
Science and technology	1,181	38.20%
Sciences	673	21.77%
Agronomical sciences and biological engineering	78	2.52%
Engineering sciences	411	13.29%
Art of building and town planning *(Polytechnique* department)	12	0.39%
Art of building and town planning (Architecture department)	7	0.23%

TABLE 13: *Descriptive statistics for the Belgian university subsample based on number of individuals, all variables (n = 1,915)*

Variable	Proportion of individuals in subsample	Final outcome in 2009 subsample %		
		Dropout	Ph.D. completion	Censored
Total	100.00%	30.60%	29.40%	40.00%
Gender				
Male	51.96%	30.35%	31.46%	38.19%
Female	48.04%	30.87%	27.17%	41.96%
Nationality				
Belgian	88.09%	29.88%	30.11%	40.01%
Foreigner	11.91%	35.96%	24.12%	39.91%
University of second cycle				
ULB	87.57%	31.01%	29.22%	39.77%
Other university	12.43%	27.73%	30.67%	41.60%
Financing				
FNRS research fellow	13.32%	10.20%	37.65%	52.16%
FRIA grant	18.22%	12.32%	43.55%	44.13%
Assistant	15.61%	28.43%	24.75%	46.82%
ULB grant	13.21%	33.99%	23.72%	42.29%
Researcher	6.11%	34.19%	35.90%	29.91%
Other type of financing	4.33%	36.14%	38.55%	25.30%
No financing	29.19%	49.37%	19.14%	31.48%
Field of study				
Human sciences	40.21%	35.71%	21.82%	42.47%
Health sciences	21.41%	32.30%	30.24%	37.56%
Science and technology	38.38%	24.35%	36.87%	38.78%
Age				
Start of Ph.D. prior to age 26	59.37%	25.59%	31.66%	42.74%
Start of Ph.D. at age 26+	40.63%	37.92%	26.09%	35.99%
Marital status				
Single	92.01%	30.19%	29.00%	40.81%
Married	7.99%	35.29%	33.99%	30.72%
Result obtained in second cycle studies				
Success without honours	4.18%	48.75%	18.75%	32.50%
Cum laude	29.14%	41.40%	22.04%	36.56%
Magna cum laude	50.23%	27.34%	31.60%	41.06%
Summa cum laude	16.45%	16.83%	38.41%	44.76%
Equivalence between discipline of second cycle and of Ph.D.				
Equivalence	91.07%	29.99%	29.99%	40.02%
No equivalence	8.93%	36.84%	23.39%	39.77%
Failed (a) year(s) during second cycle studies				
Did not fail a year	87.26%	28.31%	31.12%	40.57%
Failed at least one year	12.74%	46.31%	17.62%	36.07%

TABLE 13 (continued): *Descriptive statistics for the Belgian university subsample based on number of individuals, all variables (n = 1,915)*

Variable	Proportion of individuals in subsample	Final outcome in 2009 subsample %		
		Dropout	Ph.D. completion	Censored
Year of start of Ph.D.				
2001-2002	10.03%	42.71%	55.73%	1.56%
2002-2003	11.54%	35.75%	59.28%	4.98%
2003-2004	12.53%	42.08%	50.00%	7.92%
2004-2005	9.97%	32.46%	47.12%	20.42%
2005-2006	10.91%	33.97%	35.41%	30.62%
2006-2007	12.43%	27.73%	15.55%	56.72%
2007-2008	12.58%	28.22%	1.24%	70.54%
2008-2009	8.98%	19.77%	0.58%	79.65%
2009-2010	11.02%	10.90%	0.00%	89.10%

Notes:

26 is the median age at start of Ph.D. in the sample.

Financing and marital status are time-varying variables, which means that they can take on different values over time for the same individual. For these descriptive statistics, one value was selected per individual. For type of financing, the value selected was the type of financing that was most received during the period of doctoral study. If a student received two or more types of financing during an equal number of years, one of these types of financing was selected based on the following order: FNRS research fellow, FRIA grant, ULB grant, assistant, researcher, other type of financing and no financing. For marital status, the value selected was the one that the individual had in his last year of registration.